

An agenda for the future?

Mr Mikhail Gorbachev's speech in New York last week could be the occasion of an historic change in world affairs. But it is also a challenge for his listeners.

How does one recognize a turning-point in history? The simplest way is to wait until the textbooks have been written, but then it is too late to influence events. That is why it is an unacceptably unadventurous reaction to Mr Mikhail Gorbachev's tragically truncated visit to New York last week to decide to wait and see what happens next. Or to say that the Soviet decision to remove 10,000 tanks and 500,000 troops from Eastern Europe is a step, but too small a step, in the right direction, as did the council of the North Atlantic Treaty Organization (NATO) on Thursday last week. Gorbachev offered, within the limits of his power, to change the world. Would it not be prudent to help him succeed?

The key passage in Gorbachev's speech to the United Nations is the declaration that, because it has become "obvious" that the threat of force can no longer be an instrument of foreign policy, it follows that international relations must be "de-ideologized". "We are not abandoning our convictions, our philosophy or traditions, nor do we urge anybody to abandon theirs. But neither do we intend to be hemmed in by our values...". What this amounts to is a direct repudiation of the old Leninist doctrine of international revolution, traditionally the most powerful cause of others' suspicions of the Soviet Union's motives in international relations. Gorbachev is pragmatist enough to know the practical benefits to be won from such a posture: "the exchange of everything original that each nation has independently created". But he is also an idealist, so much so that he was described by the *New York Times* last week as "naïve". But that he means what he says is clear from the recent settlement of regional conflicts on the old revolutionary frontiers in Afghanistan and Angola. It is a pity that Gorbachev's visit to Cuba, planned for last week, had to be postponed.

The West has been caught napping by this dramatic change of course, and is calculating its reaction by the rules that have become familiar in the past forty years (since Winston Churchill's speech in 1947 at Fulton, Missouri). Are the reductions of military forces now promised sufficient to restore parity in Central Europe? And will they materialize? More generally, is Gorbachev serious? Will his colleagues let him persist with his declared policy? What, in any case, of human rights in the Soviet Union? Nobody suggests that these questions should be artificially suppressed, for they all refer to matters in which the West has a vital and legitimate interest. But the urgent need is to respond to Gorbachev's initiative in a way that demonstrates that radical and lasting change would be welcomed. Ideally, the response should be one that strengthens Gorbachev's arm, for that is a way of helping to make prophecies come true.

Falling short

NATO's response last week fell too far short of what could and should have been said, which does not imply that its anxiety about the military balance in Europe is inappropriate. Until last week, hopes for further arms control in Central Europe turned on the talks on conventional forces due to begin later this year. NATO, correctly, has been open about its negotiating position: to reduce the numbers of battle-tanks on each side to 20,000, which would require the Soviet Union and its allies in the Warsaw Pact to dispense with 20,000 tanks (twice the number

offered by Gorbachev) while leaving NATO forces essentially unchanged. There are further proposals about field artillery while NATO says nothing publicly about strike aircraft (on which the Soviet Union will consider itself at a disadvantage).

That NATO should hold to this position is entirely correct, but insufficient. The assumption is that negotiations planned more than a year ago are the only way forward. Why not also raise the possibility that, if the circumstances have or can be permanently changed, quite different ways of ensuring the security of Central Europe may be preferable. Strict demilitarization is probably impractical, although the dispersion of potentially offensive weapons to more distant places would be worthwhile. But why not test some of the outrageous possibilities that Gorbachev's statement suggests may one day be possible — that, say, of guarding against surprise attack by giving each side access to the communications traffic of the other side's military units? Or why not offer a joint study of the military arrangements that will be needed in Central Europe if Gorbachev's "non-violent world" can eventually be fashioned? This is not just an academic matter, for the opinions of Eastern European governments other than the Soviet Union would be crucial. But that, at least, would show that NATO takes him seriously.

Meanwhile, the United States, which has cautiously welcomed Gorbachev's speech, now needs to make some tangible show of its willingness to help him. The obvious way is to abandon unilaterally the Strategic Defense Initiative (SDI) — a stumbling-block in the arms control process for the past five years which is unlikely to thrive when the new US government gets to grips with the federal budget deficit after 20 January. But has not the threat of SDI been the incentive for Gorbachev's willingness to cut back on Soviet military forces? And, if SDI can be shown to function, will it not be preferable to shelter permanently behind that than to rely on vague promises of non-violence, however well-intended? These are moderate versions of the reasons often given in the United States for persisting with SDI. But they are mistaken. Gorbachev has given more convincing reasons for changing course: the need, if *perestroika* is to succeed, to release resources now spent on defence for the civilian economy. And SDI, however successful its components may be, will never be the comprehensive shield its originators believed; if pursued, it will remain a stumbling block. Why not put it on the block of mutual cooperation right away?

The other issue that has arisen in the past week is that of human rights in the Soviet Union, which has an abysmal record on these questions. How can there be an East-West understanding that there will never be another war in Central Europe when tensions constantly arise from illiberties practised in the East? This argument, valid though it may have been for most of the past forty years, may now be an obfuscation. For one thing, one of the objectives of *perestroika* is the restoration of civil rights; already there have been tangible improvements, as in the freedom of people to leave the Soviet Union. Second, it must be clear that in an issue touching the whole of civil administration in the Soviet Union, not to mention a tradition going back to tsarist times that people are the property of the state, it must take time to engineer a workable balance between personal liberty and



social obligation. That Gorbachev is at least trying should be apparent from recent events in the Baltic republics, where popular movements apparently aimed at outright secession (not generally regarded as an essential human right, especially in the United States) have been dealt with by compromise, not suppression.

Especially because, on this occasion, daring need not be compromised by military weakness, there is therefore the strongest case for taking Gorbachev's agenda for the future seriously. That his account of the future is unfamiliar to at least two-thirds of those now alive makes it seem strange, and dangerous for that reason. But that sense, that the proffered future is weirdly unfamiliar, is but another way of recognizing historic change. Will it not be cowardly to let it slip? □

Tongues shall not wag

British plans for a law on secrecy are an improvement, but are still deficient.

MOST governments have secrets. The British government has more than most, but no effective legislation to protect them. Instead, there is only the Official Secrets Act (1911), whose second (of two) sections prohibits everybody from disclosing or passing-on official information whose disclosure has not been "authorized". The result is that ministers (by definition the embodiments of authority) can selectively release public information to put a partial gloss on their own doings, but that all others, public servants and journalists in particular, do not know where they stand when using unauthorized official information. That part of the act (the other makes espionage a crime) has become unworkable: lawyers cannot easily make a case, and juries are reluctant to convict. That is why the British government is now embarked on amending legislation. A few months ago, it advertised its intentions (see *Nature* 334, 1; 7 July 1988). Now, it is pushing an amending bill through the British parliament.

These events adequately explain why previous governments have shrunk from the task. The new bill is a step forward in excluding from criminal sanction information about much routine government business. Instead, there will be six categories of protected information, dealing with intelligence, defence and foreign relations; interested parties will now know better where they stand. But the government has still not come to grips with the philosophical difficulty underlying the protection of secrets in a modern democracy, that of telling when the goal of protecting secrets, even those justly classified as such, must be overridden by other considerations, that of public discussion of intelligence activity which is illegal, or of defence procurement that wastes public money, for example.

Indirectly, the government is seeking to turn part of this objection with a second bill whose effect is to put the British intelligence services on a statutory basis, which amounts to acknowledging that they exist. Under these arrangements, the work of the services will be overseen by a committee of the great and the good reporting to the government. The same bill provides for a procedure by which the spooks can both tap people's telephones and — a new departure — burgle people's houses (court orders will be required). The intention is that the existence of the committee should meet the objection that the duty of absolute secrecy now placed on intelligence officers would stifle complaints from those who are ordered to act illegally. Sadly, in British circumstances, nobody can be sure that an internal committee will have that effect.

More generally, the government is resisting the notion that there can be a legitimate public interest in the disclosure of information otherwise classified as secret. While the bill now being debated is an improvement on the version advertised earlier in the year, particularly in its acknowledgement that those publishing secret information that should not have been

disclosed to them will be able to plead that the public interest mitigates the harm they have done, that is not the same as a full-throated acknowledgement that occasions may arise when public servants have a duty of conscience to make official secrets public. What, say, if an official knew that some future government was supplying fissile materials to other governments in contravention of its treaty obligations, or using sanctioned burglary to outwit political opponents? The Nuremberg trials after the Second World War were predicated on the principle that public servants must, in the last resort, put conscience above their instructions from above. That is the principle that should now be written into the new British legislation. □

Guns out on butter

Wealthy manufacturing nations should not also hog the world's trade in agriculture.

THE collapse last week of the trade talks at Montreal is a bad business from which everybody stands to suffer. What happened is easily described, but less easily understood. Briefly, the members of the General Agreement on Tariffs and Trade (GATT), which include the major economies of the world except the Soviet Union and those of Eastern Europe, have undertaken to negotiate among themselves an extension of the present rules governing international trade in manufactured goods to novel fields, notably agriculture and services (insurance for example). At the same time, they had undertaken among themselves to liberalize trade in the products of developing countries, from tropical products to textiles. The meeting at Montreal was meant to speed a process that has languished for too long since its beginning at a conference in Uruguay two years ago. Instead, the meeting seems merely to have confirmed the chief negotiators in their obduracy.

The central issue, on which the meeting foundered, is that of how agricultural subsidies should in future be regulated in the United States and the European Community, each of which spends roughly \$30,000 million a year at present. Logically, the extension of GATT rules to agriculture would require that all farmers' subsidies should be abolished. The obvious snag is that agricultural subsidy is a central feature of the Treaty of Rome from which the European Community derives its existence. The original idea was the mistaken notion that, in a community whose wealth stems largely from manufacturing industry, rural communities would indefinitely be impoverished if they were not offered prices for their produce linked with the general increase of prosperity. In the event, of course, agricultural prices have been too attractive, European farmers have most often over-produced — and European consumers have been overcharged for the food they eat. But agricultural protection remains central to the Treaty of Rome.

The United States, generally the cheapest food-producer in the world, has been repeatedly affronted by this state of affairs. European markets have been effectively closed (by tariffs) to low-cost US producers, while bargain sales of European food surpluses to third countries (grain to Egypt, butter to the Soviet Union) have prevented even efficient US farmers from balancing their books. Whence President Reagan's declaration two years ago that the world would be a happier place if agricultural subsidies were totally abolished by the year 2000.

The immediate cause of trouble last week was US insistence on not budging from its demand that Europe should commit itself to abolish subsidies by 2000, which would plainly have been impossible without endless prior political negotiation in Europe. If the United States had been a little more subtle, it would have been made plain that Europe is equally inflexible on several issues, not least its unwillingness to see imports replace its own farmers' produce. That is why Europe has a bigger task ahead than the United States, and — if the talks planned for the next four months should fail — will bear a greater responsibility. □

US defence production due for overhaul

- DoE safety performance criticized
- Environment clean-up on the way

Washington

A MAJOR overhaul of the US defence production facilities appears to be in the works, prompted by a series of revelations about problems at the facilities that have crippled the country's capacity to produce material needed for the production of nuclear weapons.

The manufacture of nuclear weapons is managed by a civilian agency, the US Department of Energy (DoE). Following the accident at the Chernobyl nuclear reactor in the Soviet Union, there was heightened interest in the safety of US nuclear reactor facilities. A critical report by the National Academy of Sciences in 1987 began what has become an avalanche of criticism for the way DoE manages the production of nuclear weapons (see *Nature* 334, 558; 1988 and 335, 662; 1988).

In response, DoE has begun a series of self examinations. The latest to be publicly aired is an inch-thick document detailing a myriad of environmental hazards at 16 production facilities scattered around the country. The problems, ranked using a computer model that evaluates potential risk to the population, run the gamut from groundwater contamination from the discharge of industrial solvents to soil contamination from the runoff of radioactive material. DoE estimates that the cost of cleaning up the contamination could run as high as \$70,000 million.

Heightened safety concerns also pose immediate problems for the supply of nuclear materials for weapons. All three heavy-water reactors at the Savannah River Plant in South Carolina where tritium is produced have been shut down. Last week, DoE released its restart strategy for the K reactor, initially scheduled for early next year, but now not likely to occur until next summer. The strategy acknowledges that "national security needs" require that operations start before all desirable safety modifications can be implemented. This will not go down well with an already critical Congress, but there are few options, short of borrowing some needed tritium from an ally.

Safety problems have also forced DoE to shut down its Rocky Flats Plant near Denver, Colorado, where plutonium is processed. The repairs needed to reopen the plant are not expected to be completed until February.

Two yet to be released reports, copies of which were obtained by the *Washington Post*, reportedly propose major changes in DoE production facilities. One report contains plans for spending \$50,000 million over the next twenty years to build five new production reactors to make tritium, including a heavy-water reactor at the Savannah River Plant and four high-temperature, gas-cooled reactors to be built at the Idaho National Engineering Laboratory.

This is apparently at odds with earlier reports that proposed only one new plant in Idaho. The *Washington Post* also reports that DoE is seeking \$34,000 million to \$66,000 million over the next forty years for environmental clean-up operations.

DoE could receive a boost for its Savannah River Plant restart plans this week when the Advisory Committee on Nuclear Facility Safety meets to discuss the proposal. This independent advisory group, chaired by John Ahearne, will play a crucial role in DoE's plans for restart. By publicly airing its plans to the Ahearne committee, the agency hopes to ensure a new era of openness in the way it carries out its mission.

Also meeting this week is a National Academy of Sciences committee formed to oversee the DoE nuclear weapons complex. The Academy committee is chaired by Richard Meserve, who also chaired the earlier reports critical of DoE safety performance.

Although most agree that there has been a considerable improvement in the level of concern over safety at DoE, Congress has created a new agency to make sure that any improvements achieved are not temporary. The Defense Nuclear Facilities Safety Board will consist of five commissioners appointed by the President, with an annual budget of \$7 million and a staff of 100 persons. President-elect George Bush has until March to appoint the commissioners. The Board may only recommend actions to the Secretary of Energy, but the recommendations will be published in the Federal Register, and the Secretary will have to publish his response.

A new era of safety and environmental consciousness may indeed be dawning at DoE defence production facilities, but the cost of paying for past sins will be expensive, both financially and politically.

Joseph Palca

First US patent for superconductors

Boston

THE US Patent and Trademark Office indicated late last month that it will issue the nation's first patent in the field of high-temperature superconductivity to a new process for creating a superconducting material.

The process, developed by John B. Vander Sande and Gregory J. Yurek, researchers at the Massachusetts Institute of Technology (MIT) Department of Materials Science and Engineering, yields a more malleable material than any developed previously, and overcomes the inherent brittleness of other high-temperature superconductor materials.

Using an oxidation technique, the process combines the metallic parts of a ceramic superconductor with a noble metal, such as silver. When heated and reacted with oxygen, the process converts the alloy to a composite of superconducting oxide and the noble metal. According to Vander Sande, the technique "marries the electrical properties of ceramics with the mechanical properties of metal", which leads, he says, to a material that could be formed more readily into films, wires, ribbons or coils.

Christina H. Jansen, an officer at MIT's Technology Licensing Office, says that the patent office has issued a "notice of allowability" for the patent on this process, the last formal hurdle to be faced before the patent itself is awarded. The patent is expected within the next two months.

The patent, which will be held by MIT, is already licensed to the American Superconductor Corporation, of Cambridge, Massachusetts, a small, start-up company founded in part by Vander Sande and Yurek.

Seth Shulman

Platypus egg find may be revealing

Sydney

DESPITE 200 years of scientific interest, little is known about the development or habitat of the Australian platypus (*Ornithorhynchus anatinus*), but a discovery by a research team from the University of New South Wales may shed light on the topic.

The research team came upon a female platypus's burrow containing an egg with a fetus about mid-way through its incubation period. The fetus is being dissected at the University of Queensland.

The find is said to be the first example found since the 1880s of a fetus at such a stage of development. There has been only one instance of a platypus breeding in captivity, and that was 45 years ago. Platypuses are monotremes, a subclass of mammals which lay eggs but suckle their young.

Tania Ewing

Aftermath of earthquake

Armenians face up to scenes of total devastation

Moscow

THE cataclysm that struck Armenia on 7 December claims as of 12 December 40,000 to 45,000 lives and tens of thousands wounded. Over half a million people have lost their homes. The town of Spitak with a population of 20,000, at the epicentre of the worst earthquake in Armenia in the past thousand years, was levelled.

Eighty per cent of the homes and industrial structures were reduced to rubble in Leninakan, Armenia's second largest city (population: 290,000), and half of all structures in Kirovakan (population: 170,000). The earthquake hit an area with a population of more than 700,000, toppling buildings and severing power, gas, telephone and other communication lines. These are preliminary estimates of the Armenian calamity.

President Mikhail Gorbachev cut short his visit in New York, postponed planned visits to Cuba and Britain and arrived in Armenia to take charge of the situation. Prime Minister Nikolai Ryzhkov has been in Yerevan right from the start, coordinating the rescue and salvage operations as head of a special CPSU (Soviet Communist Party) Politburo commission.

Other members of the commission working in Yerevan, the capital of

ters. The report was based on seismic records of the tremor and on macroseismic occurrences in Tbilisi and Bakuriani, Georgia. Within the next thirty minutes, additional information arrived from twelve seismological stations. More accurate parameters of the calamity could not be had directly from the disaster area because the seismological station in Leninakan had been cut off.

The civil defence headquarters, the CPSU Central Committee and army units were informed of the earthquake just before midday, and the Institute of the Physics of the Earth of the Soviet Academy of Sciences, the coordinator of the country's seismological service, at 12:02 p.m. Urgent reports continued coming in till 1:40 p.m. from 58 seismological stations.

From the incoming data, the following parameters were established: τ (time) = 7:41:23.1 GMT; φ = 40.82 degrees lat. N; λ = 44.05 degrees long. E; h (depth) = 3 kilometres; M (magnitude) = 7.0.

Deciphering the data shows that the earthquake began three kilometres beneath the surface with the epicentre at Spitak, between Leninakan and Kirovakan. The shallowness of the earthquake lent it strength and intensity.

From the signs of the P-waves, seismologists have been able to reconstruct the possible earthquake mechanism. There are two alternative fault planes, specified by the following parameters:

	Strike	Dip	Slip
NP1:	276°	15°	147°
NP2:	38°	82°	78°

The same analysis gives the directions of maximum tension and of compression as azimuth 294° (plunge 51°) and 139° (plunge 36°) respectively. The intermediate direction has azimuth 40° (plunge 12°).

Seismological stations also recorded the following macroseismic tremors on the 12-point Soviet scale: 9 points at Leninakan; 6 to 7 points at Bakuriani, Tabatskuri and Borzhomi; 5 to 6 points at Bogdanovka, Tbilisi and Yerevan; 5 points at Goris; 4 points at Makhachkala; 3 to 4 points at Grozny; and 3 points at Sheki and Shemakha.

Igor Nersesov, head of the Institute of the Physics of the Earth, says that the underlying cause of the earthquake is that "The Arabian plate is moving towards the region from the south, and the accumulated pressure of the movement is released in earthquakes."

There have in the past been powerful earthquakes on the border between Armenia and Turkey, especially between AD 900 and 1000. One of the Armenian seismological stations is situated near Aini, the ancient capital of Armenia



Relief workers in Leninakan.

toppled by the earthquake in AD 1046 of 8 to 9 points. The last big earthquake, magnitude 5.7 (about 8 points) was in Leninakan in 1926. But in intensity and the devastation caused, last week's earthquake is comparable with those in Tangshan, China, in 1976 and Mexico City in 1985.

Last week's events were not predicted. But earthquakes usually occur in regions of intense seismic activity and seismic zones are marked on the seismological maps which are used in the Soviet Union, as elsewhere, in architectural design for earthquake-prone areas as well as for earthquake-prediction and risk-reduction measures.

The Politburo commission is still working at Yerevan, but Soviet experts are already very critical of the flawed design and construction which added to the earthquake's heavy toll in Leninakan. Recommendations of scientists and engineers seem to have been largely ignored. It is significant that multi-storey structures built only three years ago crumbled, while the lower houses built twenty and more years ago remained standing.

On the question of whether it would have been possible to have predicted the earthquake, no detailed information is as yet available, for telephone lines to the Leninakan seismological station had not been restored by 11 December.

"We do not have the necessary technical facilities for prompt data processing, which is essential for sufficiently accurate forecasting of earthquakes," says Nersesov. "If we had had such facilities, we could probably have predicted the oncoming catastrophe in Armenia and reduce its



Area affected by last week's earthquake.

Armenia, include the ministers of defence and health and several high-ranking officials. Large numbers of experts, plane-fuls of special equipment, blood, plasma, medicine, food, tents, warm clothing and other prime necessities have been coming to Armenia from all over the Soviet Union and more than 40 foreign countries.

The earthquake hit the area on 7 December at 10:41 a.m. Moscow time. Within the next thirty minutes, nine seismological stations in Kizyl-Arvat, Tbilisi, Yel'tsovka, Yakutsk, Uzhgorod, Moscow, Bakuriani, Tashkent and Ashkhabad sent emergency reports to the central seismic observatory at Obninsk, near Moscow.

Thirty minutes later the first urgent report of a powerful earthquake in Armenia reached the Council of Minis-

toll."

The leading Soviet seismologist also calls into question the expediency of building nuclear power stations in earthquake-prone regions, even if their design is technologically up to standard. The Armenian nuclear power station designed to withstand tremors of up to 8 points did not suffer in the earthquake, but there was a real risk of damage.

One striking feature of last week's catastrophe is that the Soviet press has been full of information about the disaster. This, a consequence of the leadership's policy of *glasnost*, is in sharp contrast with the reticence following the disasters in Ashkhabad in 1948 and Tashkent in 1966. One of the side benefits has been that the relief effort has assumed unprecedented proportions.

Meanwhile, the earthquake has made people forget their differences and join forces in grief, compassion and assistance.

Vladimir Markov
Novosti

Vera Rich writes: Markov is certainly correct to describe the handling of the Armenian earthquake as a telling example of how far Soviet *glasnost* has progressed, even since the Chernobyl disaster of April 1986. Then it was more than 60 hours before an official announcement was made and, even then, the accident was treated as a minor news item. Last week, however, the leading Soviet papers carried at least a brief report in the first available edition, and reversed a long-standing Soviet tradition by carrying disaster coverage on the first page.

The Soviet press also reports that the bureau of the Central Committee of the Communist Party of Armenia was in session when the earthquake struck. The first reaction of the bureau members was to telephone to the Armenian nuclear power station at Medzamor, some 35 km from Yerevan, as well as to the major chemical enterprises in Armenia.

But so far it has not been remarked that both the nuclear power station and the chemical plants have been during the past three years the subject of an environmental campaign backed by several members of the Armenian Academy of Sciences, which particularly criticized the siting off such establishments in a seismic zone. A few weeks ago, in response to this pressure, it was promised that the Medzamor power station would be phased out by 1991, the earliest date by which alternative energy supplies could be ensured.

Similar concerns have more recently arisen in Tadjikistan, the earthquake-prone mountainous republic further east, where it is claimed that hydroelectric dams have been built using safety factors worked out for single dams in low-lying aseismic areas. □

Conflicting views on extent of earthquake threat to Tokyo

Tokyo

A REPETITION of the Great Kanto Earthquake of 1923 would result in widespread death and destruction in the Tokyo area, according to a report released last week by the National Land Agency. The report has shocked Tokyo residents, but critics say it fails to gauge the full possible extent of the disaster.

Tokyo lies in one of the most earthquake-prone regions of the world — second in the list of fast-growing cities 'at risk', on page 626 of this issue. To the east, the Pacific plate plunges beneath Japan and to the south-west the subducting Philippine plate slices into the Izu Peninsula between Suruga and Sagami bays. In 1923, the Great Kanto earthquake, which was centred in Sagami Bay and registered 7.9 on the Richter scale, destroyed large sections of the city and left nearly 150,000 people dead.

To assess the impact of a similar event, the agency divided Tokyo and surrounding prefectures (Saitama, Kanagawa and Chiba) into a 1- or 2-kilometre grid and modelled the effects of the 1923 earthquake on present-day Tokyo, taking into account conditions in each block of the grid. The results suggest a similar death toll to the 1923 event.

Modern buildings within Tokyo are designed to withstand major earthquakes (see below). But in many parts of the city, old wooden houses can still be found squashed together within inches of each other and conditions are ripe for disastrous fires.

If an earthquake should strike in the middle of the night, the death toll would be limited to 80,000, according to the report. But if it came at lunchtime (as in 1923) or early in the evening in winter when stoves would be lit for cooking or heating, fires could be expected to sweep through parts of the city soon after the earthquake and the number of deaths could reach 150,000, with another 200,000 injured, the report says.

But some critics say the report may underestimate the full extent of the damage that might be caused by a major earthquake. For example, no account is taken of recent reclamation and development projects along Tokyo Bay waterfront and no estimate is made of the possible death toll resulting from panic in the underground railways and shopping malls scattered around modern-day Tokyo.

Although computers have back-up systems to cover such events, some economists fear that a major earthquake in Tokyo could cause panic on the world market.

The National Land Agency report, which also maps out responsibilities for the various government ministries and agencies in the event of such a disaster, has been submitted to the Central Disaster Prevention Council, chaired by Prime Minister Noboru Takeshita, which will propose counter measures.

But the only effective solution would probably be decentralization and drastic changes in the planning and development of the city.

David Swinbanks

Builders look to 'anti-quake' device

JAPANESE engineers have developed an ingenious way to protect tall buildings from earthquakes. A device placed on the top of the building creates 'anti-quakes' to cancel out the effects of earthquake motion.

Tall buildings in Tokyo and other major Japanese cities are designed to flex and sway during earthquakes. The oscillations absorb the earthquake and protect the building from collapse. But the swaying motions can be a frightening experience for people on the top floors even during comparatively minor earthquakes.

Kajima Corporation, a construction company, plans to bring the swinging to a halt with 'anti-quakes'. Kajima's device consists of huge weights (several tons) on wheels placed on the roof of the building. A computer connected to motion and vibration sensors in the basement and mid-floor of the building command hydraulic actuators to move the weights back and forth very rapidly in such a fashion as to cancel

out the forces of the earthquake.

Kajima researchers say their device can reduce motions up to 75 per cent in earthquakes up to magnitude 4 on the Japanese scale of 7. Magnitude 4 earthquakes cause buildings to shake violently, knocking over ornaments and glasses. The device is also effective in preventing motions induced by strong winds of up to 70 kilometres per hour.

By May next year, Kajima Corporation will install its first commercial anti-quake device in an 11-storey building now being built in Kyobashi, Tokyo. The device, however, does suffer from one potential hazard. If the timing is slightly out, the moving weights will amplify rather than decrease oscillation of the building. But Mitsuo Sakamoto of Kajima's Kōbori Research Institute says that the computer software contains a "fail-safe" program that can switch the system off if things get out of control.

David Swinbanks

CDE faces verification problems for chemical weapons

Porton Down, Wiltshire

SCIENTISTS from the UK Chemical Defence Establishment (CDE) who exchanged visits this year with staff from the equivalent Soviet facility at Shikhany, 600 km south-east of Moscow, were neither given access to a set of buildings that are clearly identifiable on satellite photographs, nor provided with satisfactory information about them, despite requests. They suspect the main building is a chemical production plant of some kind. In general, says CDE's director, Dr Graham Pearson, the Soviets have revealed very little of their developments in chemical weaponry over the past 30 years.

A major cause of concern is the possible development of chemicals that are closely related to biological agents but may not be covered by the ban imposed by the biological weapons convention of 1972. The blurring of the distinction between chemi-

cal and biological weapons is exacerbated by recombinant DNA technology. This would allow the production in quantity of versions of bacterial toxins and potentially toxic neuropeptides, such as Substance P, that were sufficiently altered arguably to be classified as chemicals.



Cheers: the new "General Service Respirator" from Porton is the first to allow the wearer to drink safely in a contaminated environment.

cal and biological weapons is exacerbated by recombinant DNA technology. This would allow the production in quantity of versions of bacterial toxins and potentially toxic neuropeptides, such as Substance P, that were sufficiently altered arguably to be classified as chemicals.

The possible production and deployment of new chemical agents is adding to the research load of the 400 staff at CDE, the Ministry of Defence establishment that claims to work solely on defensive and protective measures against chemical and biological weapons. Underlying their concern is the fact that many of the newer chemicals are so closely related to substances with a genuine industrial use that it would be easy to build a dual-production facility and disguise its second purpose.

That possibility is exercising the minds of those concerned with verification measures. A second verification problem is the identification of new or even classical chemical weapons in field

samples. There is little doubt from firsthand descriptions that chemical weapons were used in Cambodia, says Pearson, but their identity is unknown.

For detection of chemical agents in the battlefield, CDE is developing second-generation ion mobility spectrometers. The most promising has two drift tubes sharing a common source region, which allows the simultaneous detection of positive and negative ions, but it will need considerable software development to enable such spectrometers to detect and give warning of the presence of any chemical weapon.

Other current research at CDE includes the improvement, by chemical modification, of the adsorption properties of the granular charcoal that is the main protective substance in gas masks and anti-gas

fabric, and the development, in conjunction with Roche, of a soluble pro-drug form of diazepam, which combats the convulsive effects of nerve gases. The pro-drug will be added to the other two soluble agents in the self-administration syringe currently supplied to UK forces along with a diazepam tablet. Another industrial collaborator is Celltech, the UK biotechnology company. Using the company's know-how in monoclonal-antibody-based diagnostics, CDE biologists have developed a 10-minute field-test for *Clostridium perfringens* phospholipase C, the major toxin responsible for gas gangrene. They are also working on an improved vaccine against the bacterium.

There is a danger that the production of effective protection against existing chemical and biological weapons will encourage the development of new weapons, admits Pearson, which is why CDE "does not publish precise performance parameters of equipment".

Peter Newmark

Alvey programme under fire

London

THE UK government's information technology programme, financed jointly with private industry and intended to rectify the serious adverse balance of trade in information technology products, was criticized last week in a report by the House of Commons Committee of Public Accounts. The £350-million five-year Alvey programme, which began in 1983, suffered from understaffing and poor management, says the committee. Because three government departments were involved, there were problems of contracting, financial control and financial support.

Severe delays were caused because project participants found it difficult to secure collaboration; the Alvey directorate could have done more to help, says the committee. It also criticizes the decision to direct more money than planned into research in academic institutions, arguing that savings from carrying out research more cheaply in universities are outweighed by the loss of funds from industry.

Although the committee praises the directorate for instigating cooperation between academic institutions and industry, it says that there was much less participation than planned. And it says the project was affected by the country's skills shortages in information technology.

The first five years of the project were to have been followed by a second phase. But the programme will not now be continued; in future, the country's main information technology effort will be channelled through the European ESPRIT programme.

Christine McGourty

Lecturer resigns over MOD funding

London

A GRANT of £235,000 from the Ministry of Defence (MOD) to the veterinary school of the University of Bristol has led to the resignation of at least one academic member of staff on the grounds that the work has links with biological warfare, although this is denied by the university.

The project at the centre of the debate is to be carried out in the department of animal husbandry and will concentrate on the study of the organism, *Klebsiella pneumoniae*, which those protesting against the grant say causes mortality in man, but is of no significance to animal respiratory disease. Dr Sue Mayer, who has resigned from her position as lecturer, says she cannot work there because she feels the improvement of animal health and welfare is a 'spin-off' rather than the principal justification of the work. Another member of staff is also said to have resigned.

But the grant-holder, the head of the department and the vice-chancellor of the university have no complaints about the funding of the project by the MOD and deny charges that it was redesigned to suit the MOD after the Agriculture and Food Research Council refused to support it.

The Association of University Teachers, which originally spoke out in support of those protesting at the grant, has rescinded that support and apologized to the university.

The start of the project has now been delayed; a spokesman for the university says that this is due to the reluctance of the MOD to pay overheads and certain expenses.

Christine McGourty

German cabinet promises gene law but problems remain

Munich

GIVING birth to a new 'basic law' for regulating genetic engineering is giving the West German government labour pains. A Health Ministry report on the outline of the new law was received by the cabinet with a great fanfare in Bonn last month, but it is clear that important issues remain to be decided if the law is to be enacted before the end of next year.

The government hopes that a single, general law will close legal loopholes in the present patchwork of regulations. At the same time, the law will not be too specific. Instead, it will ensure that the regulations are regularly reviewed and updated to keep up with the technology.

The new law will probably not change regulatory procedures for universities that are carrying out basic research in

biotechnology.

Applications to conduct projects will be submitted to a central body such as the existing Central Commission for Biological Safety (ZKBS), which operates under the auspices of the Federal Health Office in Berlin. ZKBS can at present only request voluntary compliance with its decisions, whereas the new law could make violating a ZKBS order a criminal offence.

For industry, the situation is more complicated, especially in the case of intentional release of genetically modified organisms into the environment. The new law may closely resemble the pattern established in pollution control laws.

The extent to which the new law will mesh with or take precedence over existing laws is still undecided. The Health Ministry report urges that individual ministries give up much of their authority in areas they regulate to make the new law as broadly applicable as possible, but some, especially research and technology, agriculture and environment, may well resist.

The question of whether to declare a five-year moratorium on any release of a genetically modified organism is still open, to the exasperation of researchers. Such a moratorium was recommended by a parliamentary commission in January 1987. Nearly two years have gone by and a five-year moratorium is still a possibility.

The federal government gives itself precedence over the *Länder* in enforcing the new law, a move which will require the approval of the *Länder* and might provoke a political battle.

Liability law will certainly change. At present, West German law requires a clear chain of causality to be demonstrated in establishing liability. But in the case of a lawsuit against a university or industrial research institution for damages caused by genetic engineering, the government will make exceptions.

Areas such as genome analysis, gene therapies and the protection of human embryos in research were explicitly excluded from the new law, since the Justice Ministry is already hard at work on proposed laws covering those subjects. The results are expected by midsummer.

Basic researchers and West German industry are anxious for the law to be passed. Industry has postponed investments in the area of genetic engineering in anticipation. The government has promised to pass the law before the next parliamentary elections in late 1990, but it must resolve the remaining issues by the midsummer recess in 1989 to allow for parliamentary debate. **Steven Dickman**

Human gene therapy now on the way

Washington

THE first experiment involving the introduction of foreign genes into humans got back on track through the institutional review process last week when it was approved by a subcommittee of the Recombinant DNA Advisory Committee (RAC) of the US National Institutes of Health (NIH).

The project, directed by NIH researchers Stephen A. Rosenberg and W. French Anderson, became derailed last month when NIH director James Wyngaarden required that the subcommittee reevaluate the project because they had not been given adequate information initially. This despite the fact that the full RAC had already approved the project (see *Nature* 335, 577 and 754; 1988).

The protocol calls for treating ten cancer patients by giving them tumour-infiltrating lymphocytes (TIL) marked with genes for antibiotic resistance. The genes will have no therapeutic value to the patients, but will allow the cancer-killing TIL cells to be tracked in the patients' bodies.

The NIH institutional biosafety committee cleared the experiment early last month, but cautioned that its approval should not be construed as setting a precedent for allowing human gene therapy.

The protocol must now be approved by the National Heart Lung and Blood Institute's institutional review board, and be reapproved by the RAC. If it clears both these hurdles, the protocol will then be reviewed by the US Food and Drug Administration and NIH Director James Wyngaarden.

Carol Ezzell

Atlantis back safely

THE secret defence mission of the space shuttle Atlantis ended successfully on Wednesday afternoon, 7 December, at Edwards Air Force Base in California, four and a half days after it lifted off from the Kennedy Space Center in Florida. Although mission details are classified, the reported goal was to deploy a high-resolution radar imager capable of mapping the Earth's surface in any weather condition. Atlantis's launch trajectory indicated that the satellite will orbit over the Soviet Union.

The only reported problem Atlantis encountered was the unusually large number of heat-shield tiles that were damaged during the mission. Between 125 and 175 tiles were either partially or completely destroyed. The National Aeronautics and Space Administration (NASA) has established a board of inquiry to determine the cause of the damage.

The next shuttle to fly will be Discovery, scheduled to place another Tracking and Data Relay Satellite into orbit next February. Atlantis's next flight, set for April, will be the long-awaited launch of the Magellan radar mapper mission to Venus. **J.P.**

US award to France

ON the occasion of its centenary celebrations, the US Geological Society awarded its gold medal to Claude Allègre, professor at the University of Paris VII and long-time director of the Institut de Physique du Globe



and the laboratory of geochemistry and cosmochemistry that he founded in 1965. It is the first time the society has awarded this distinction to a foreign scientist. Allègre has pioneered studies of the evolution of the Earth based on the isotopic composition of rock samples and has been instrumental in initiating several major international projects in geochemistry and global dynamics. In 1986 he received the Crafoord Prize from the Swedish Royal Academy. **P.C.**

SERC falls behind

INCREASING inflation is causing concern within the Science and Engineering Council, and chairman of the Science Board, Professor Brian Fender, admitted that figures in the council's annual report, published last week, show that funds for research grants are not increasing as rapidly as he would like.

He also said the council was concerned that the number of postgraduate researchers might drop, and to prevent this the council was considering raising the level of grants. In addition, Fender said that the University Grants Committee reviews of physics and chemistry in universities which designated minimum sizes for departments were too inflexible; it will put its comments on the reviews to the full council this week. **C. McG.**

Australia losing more brains

Sydney

AUSTRALIA'S brain drain is intensifying, according to a new survey by a senior lecturer in mathematics at the Australian National University in Canberra. An entire generation of the country's most outstanding mathematicians seems likely to be lost to overseas institutions.

The survey, carried out by John Hutchinson, tracked 62 students who graduated with first-class honours in mathematics during 1970–87. Thirty-five of them are involved in research and development and thirty-two have received doctorates. But none is working in Australian industry and more than half have gone overseas, mostly to the United States.

Recently, the percentage going overseas has increased. Seventy-five per cent



of those who graduated in 1970–75 are still in Australia. But only 35 per cent of those who graduated in the following five years have remained at home. The report concludes that while improved salaries would help to retain and attract scientists, the major consideration is a stimulating and stable research environment.

According to Jim Peacock, chief of the Division of Plant Industry within the Commonwealth Scientific and Industrial Research Organization (CSIRO), it costs A\$500,000 to train a scientist to the end of a postdoctoral fellowship. "At the end of this period they are at their most productive for Australia and we should be doing all we can to keep them", he said.

Scientists who have remained in Australia are becoming increasingly angry about their plight. The Prime Minister, Bob Hawke, was heckled by 700 scientists at the recent opening of the new National Science and Technology Centre in Canberra. The scientists were protesting at continuing cuts in government support for scientific research. Hawke attempted to calm the crowd by promising increases in funds for scientific and technological research. A committee of senior civil servants has been set up to report on the need for further funds.

Tania Ewing

NSF providing support to get new centres off the ground

Washington

THE US National Science Foundation (NSF) last week announced that it will spend \$25 million in fiscal year 1989 on eleven Science and Technology Centers intended to foster cooperative research among university, industrial and national laboratories. The intended benefit, says NSF director Eric Bloch, is "a much shorter time-span between actual discovery and utilization".

Although the centres are intended to increase the ability of the United States to compete in world markets, it proved impossible to find funds for them last year. This year, Congress turned down a request for a separate five-year \$150-million appropriation for the centres but gave NSF enough funds for the first year of operations.

Competition for the centres was severe. NSF received 323 proposals and employed more than 1,000 reviewers to help whittle them down to a short list of 48 finalists. Site visits from a 23-member multidisciplinary panel resulted in the choice of 11 centres for awards ranging from \$900,000 to \$4 million.

The centres are university based, and with one exception focus on topics where basic research seems likely to reap financial dividends. The exception is the University of California at Berkeley which will receive \$1.8 million for a Center for Particle Astrophysics under the direction

of Bernard Sadoulet. In collaboration with other California universities and the Lawrence Berkeley Laboratory, the centre will search for the 'essence' of the dark matter, known only by its gravitational effects, that makes up 90 per cent of the mass in the Universe.

The remaining centres and their budgets are: ■ University of California at Santa Barbara Center for Quantized Electronic Structures (\$2.1 million) ■ California Institute of Technology Center for the Development of an Integrated Protein and Nucleic Acid Technology (\$3.0 million) ■ University of Illinois at Champaign-Urbana Center for High-Temperature Superconductivity (\$4.2 million) ■ Michigan State University Center in Microbial Ecology (\$1.1 million) ■ Northwestern University Center for Advanced Cement-Based Materials (\$1.7 million) ■ University of Oklahoma Center for Analysis and Prediction of Storms (\$0.9 million) ■ Rice University Center for Research on Parallel Computation (\$4.1 million) ■ University of Rochester Center for Photoinduced Charge Transfer (\$1.6 million) ■ Rutgers University Center for Discrete Mathematics and Theoretical Computer Science (\$1.8 million) ■ Virginia Polytechnic Institute and State University Center for High-Performance Polymeric Adhesives and Composites (\$2.1 million).

Alun Anderson

South African doctor sacked "inexplicably"

Oxford

A MAJOR controversy has erupted in South African medical circles over the transfer of the chief superintendent of Cape Town's Groote Schuur Hospital, Dr Jocelyn Kane-Berman, after she was quoted in a weekend newspaper as having named Nelson Mandela as her choice as prime minister in a South African cabinet selected on merit.

Kane-Berman was transferred from her post at the end of last month to be regional medical superintendent for the Western Cape. The administrator of the Cape Province, Mr Gene Louw, issued a statement on 9 December confirming that Kane-Berman's transfer was a disciplinary measure, and refusing to reinstate her. He justified this decision on the grounds that Kane-Berman's comments were "closely linked to radical politics". He added: "In the present political climate it is top priority for all provincial officials, in the execution of their responsibilities, to desist at all costs from personal political involvement or statements which may harm the provincial government service."

Condemnation of Kane-Berman's transfer has been widespread. Dr Stuart Saunders, vice-chancellor of UCT and former head of its department of medicine, said that he was "dismayed" by the administrator's decision, and called for the reinstatement of Dr Kane-Berman. The Public Servants League described her removal as "inexplicable considering the reasons so far". Even the federal council of the conservative Medical Association of South Africa (MASA) has asked for her reinstatement. Both MASA and UCT have promised to support Kane-Berman if she decides to appeal to the courts to overturn the administrator's decision.

Perhaps the most pertinent comment came from Dr Marius Bernard, the Progressive Federal Party's spokesman on health. He criticized the administrator for claiming that Kane-Berman had brought politics into the health services, adding: "The government has politicized the health service and administrators National Party policy daily, as can be seen in its segregated hospitals".

Michael Cherry

Clean bill of health for Sellafield and Dounreay reactors?

London

NEW evidence for a theory which attempts to account for the observed higher-than-average incidence of childhood leukaemia near nuclear reprocessing plants in the United Kingdom is published in the current issue of *The Lancet*. Leo Kinlen of the Cancer Research Campaign's epidemiology unit at the University of Edinburgh, suggests that the rapid influxes of population near reprocessing plants are responsible for the increased leukaemia rates, not radiation discharged from the plants, as is popularly believed.

The apparently increased incidence of childhood leukaemia around the two nuclear reprocessing plants in the country — at Sellafield and Dounreay — is no longer much disputed. But there is no accepted explanation for it.

It is well known that people living in isolated places tend to escape common infections, so that when outsiders move

in, an epidemic can occur because the indigenous population lacks resistance to the infective agent. And it has been proposed that leukaemia clusters in areas which were isolated and then experience a population influx, such as Sellafield and Dounreay, result from a virus infection to which the indigenous population has not evolved resistance.

Kinlen set out to test this theory by looking for a similarly high incidence of leukaemia mortality in an area which had been isolated before experiencing a sudden population influx, but which lacked any man-made potential source of radiation. In the area identified, Glenrothes in Scotland, Kinlen found an excess in deaths from leukaemia in the under-25 age group, particularly in the under-fives, when he compared the mortality rates in this area to the national average. Only 3.6 mortalities in the under 25 group were expected and nine were found; in the under-5 group only 1.5 were expected but 6 were found.

Kinlen says that this work is consistent with the report published in June of the Committee on Medical Aspects of Radiation in the Environment (COMARE). That committee found six cases of childhood leukaemia in the Dounreay area between 1979 and 1984, when only one was expected (see *Nature* 333, 587; 16 June, 1988). And it said that this evidence, combined with evidence of increased leukaemia incidence at Sellafield, suggested that some feature of the nuclear plants — not necessarily radiation — leads to an increased risk of leukaemia in young people living in the vicinity.

Both BNFL and the United Kingdom Atomic Energy Authority, which runs Dounreay, welcome Kinlen's work, saying that they have long disputed that radiation could be responsible for increased cases of leukaemia around the plants and that this work seems to bear them out. But Simon Boxer, of the pressure group Cumbrians Opposed to a Radioactive Environment (CORE), says that it will take a lot to convince the people living near Sellafield that radioactive discharge is not the culprit. Ray Cartwright of the Leukaemia Research Fund unit at the University of Leeds says the new research is consistent with the hypotheses that his unit is investigating to explain leukaemia clusters. But he says it seems unlikely that a single virus may be responsible; several viruses or a plethora of infectious agents could be involved. He also says there is strong evidence that low-dosage radiation from natural or man-made sources may play an important part.

Christine McGourty

No change at IAEA?

HANS Blix revealed last week that he will seek a third term as director-general of the International Atomic Energy Agency (IAEA), and it is very likely that the agency's board of governors will choose him when it meets in June. The board's selection must be approved by two-thirds of the 113 member states at the September 1989 general conference. The new term begins in 1990 and lasts four years.

Developing countries belonging to the Group of 77 earlier this year expressed dissatisfaction with Blix, but they have not been able to agree on a candidate to challenge him (see *Nature* 333, 288; 1988). Blix, a Swede, is only the third director-general in the 31-year history of the IAEA. There has been no director-general from a developing country.

IAEA is an independent body that monitors compliance with the international treaty on nuclear non-proliferation as well as providing technical assistance in nuclear safety, medicine and agriculture and in nuclear energy development. S.D.

Synchrotron city

THE world's most powerful synchrotron seems set to be located in a new science city in western Japan. The Science and Technology Agency last week gave the top ranking to Nishiharima technopolis in Hyogo Prefecture, west of Osaka, in a land survey of four potential sites for a giant synchrotron which the agency plans to build by 1995.

The new synchrotron, which will have a storage ring half a kilometre in diameter and will be able to accelerate electrons up to 8 GeV, will be the world's largest; a 6-GeV European facility is under construction in Grenoble and the United States is building a 7-GeV synchrotron at the Argonne National Laboratory in Illinois. Final cost of the machine is expected to be some ¥100,000 million (\$800 million) and the agency has requested ¥2,051 million for its development in fiscal year 1989.

There has been strong competition among local governments to host the new facility because it is expected to help development of high-tech industry, in particular the semiconductor industry. A decision will be made after the completion of additional surveys and an environmental impact assessment. D.S.

Hurd on animals

A CODE of practice on the use of laboratory animals for scientific purposes in the United Kingdom will be issued by the Home Secretary, Douglas Hurd, by the end of the year. It will provide researchers with guidelines in line with tighter controls on the use of animals introduced in the Animal (Scientific Procedures) Act 1986. The animal procedures committee, which published its first report last week, will help to develop the code of practice under consideration. The code is based on work done by the Royal Society and the Universities Federation for Animal Welfare. C. McG.

Laboratory accident kills AIDS mice

Washington

THE first experiment involving transgenic mice containing the genome of HIV, the virus causing AIDS, ended last week when the mice died from accidental overheating.

The transgenic mice were created as an animal model to mimic the symptoms of AIDS in humans. Malcolm Martin, the lead researcher in the experiment being conducted at the US National Institutes of Health (NIH), expected that mice containing HIV genes would produce HIV proteins and perhaps develop AIDS.

Martin received a phone call from a technician at 5:00 a.m. on Monday last week telling him that all but three of his 130 mice were dead. The ventilation in the laboratory housing the mice had been inadvertently cut off for at least seven hours during maintenance work over the weekend. The mice were being kept in a specially designed 30-foot laboratory glovebox to prevent them from escaping and mating with wild mice, passing on their HIV genes. The glovebox had an autoclave at each end, and the mouse cages inside the box were surrounded by a moat of bleach.

Martin says the shutoff of the glovebox ventilation caused the illumination of an alarm light at NIH's central control centre, but the light was ignored because it was thought to be malfunctioning.

No foul-play is suspected in the accident, but the NIH engineer who neglected to check the building after the warning light went off may be reprimanded. Carol Ezzell

McBride criticizes inquiry

SIR—You have reported the findings of the committee of inquiry into the allegations of scientific fraud by me (*Nature* 336, 101; 1988). I believe this is the first inquiry of its kind in the world. I hope it will never be repeated.

Allegations of scientific fraud were first made by Norman Swan, a medical reporter for the Australian Broadcasting Commission, on a radio programme, "The Science Show", on 12 December 1987; they were based on statements by two of my former assistants and were repeated in the lay media.

The board of directors of Foundation 41, where I have served as medical director for 17 years, then set up an inquiry chaired by Sir Harry Gibbs, a retired chief justice of Australia, and including two scientists nominated by the Australian Academy of Science, Professors Robert Porter and Roger Short. Neither has published on birth defects, the mechanisms of their production or fetal pharmacology.

After I had agreed to assist this inquiry, I was informed that no party was to be allowed legal representation. My lawyers did present the committee with a written request on my behalf emphasizing that I, the accused, had my reputation to lose and should have the right of representation. I was advised by my legal advisers not to participate in the inquiry if this request was denied, but as I had already agreed to participate and as I maintain that I am innocent of the charge, I foolishly decided to attend.

The inquiry was set down for 25, 26 and 27 July 1988. As my accuser, Norman Swan, was about to travel overseas, a special sitting was held for him to give evidence on 7 July.

Not only was I denied legal representation, but I was not allowed to attend the inquiry and thus question my accusers. I had to put in a written submission before 30 June because of Swan's impending departure. As I had little time to prepare it and had to submit three copies, much work was involved. I was allowed to make later additions.

The inquiry was advertised in every major newspaper in Australia and the *Medical Journal of Australia*. Anyone could make a submission. As I was present only for approximately one and a half hours on 25 July and a short time for questioning on Wednesday 27 July, I still do not know who gave evidence or what evidence was given.

The experiment in question was entitled "Effects of scopolamine hydrobromide on the development of the chick and rabbit embryo," published in the *Australian Journal of Biological Science* (38, 173; 1988).

I gather that much of the evidence against me was supplied by Mr Fred Baker, a Washington lawyer who represents the Merrell Dow Company (this company does not manufacture scopolamine). This evidence from Mr Baker consisted of affidavits that would not be accepted as evidence by a court because the people who made them were not present for questioning.

Professor Short is a frequent speaker on the Science Show. On 29 October, he spoke on AIDS. After his segment, Mr Robyn Williams, the producer, said that Short had been a member of the committee investigating the charges of scientific fraud made by Swan against me on 12 December, and that the committee's findings would be handed down on 2 November. This was the first I had heard of the date of release of the findings. Williams went on to congratulate Swan on receiving the Walkley Award for journalism for exposing Dr McBride's fraud. This award had been announced on 27 October, six days before the committee's findings were released. Short's appearance and the announcement of the Walkley award before the Gibbs Committee report was handed down seems extraordinary. My criticisms of this procedure are as follows.

1. I was denied legal representation although I was the accused and I was refused the right to cross-examine my accusers.
2. I still do not know who gave evidence against me or the nature of the evidence.
3. I have requested but not had my original documents returned. These contain correspondence, some of which is eight years old, and a manuscript of the same time. I was not warned before I submitted this material that all three sets of the documents would be retained by the committee. Sir Harry Gibbs has ruled that this material is now the property of the committee.
4. Although the recorded transcript of the proceedings was paid for by Foundation 41, this has not been given to the board of directors.
5. I find it impossible to accept that scientists would dismiss eight malformed rabbit kittens, all with rare malformations of the brain or eye or both, as probably due to disease or age of the mother. It was suggested by Short on 25 July that rabbit food could have caused the deformities, because cat food can be toxic to cats. I did a computer literature search on 26 July and reported to the committee on 27 July

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

that I could find no evidence of malformations produced by food. Even the Minamatta Bay disaster (mercury contamination of fish) did not produce physical malformations.

6. I feel the evidence collected by the committee may have been influenced by the partisan attitudes well described by the *Sydney Morning Herald* correspondent who, in a report of the inquiry, said of me that "he was a doctor, working in another professional area — medical science — where he was never quite accepted". Another example is the casual way in which the date of its finding was first announced, as an addendum to a popular broadcast.

As this inquiry sets a precedent, I hope that no one will be subjected to a similar procedure. Unfortunately, there is no way that I can appeal against the decision. I still maintain I am innocent of the charge of scientific fraud. Other workers have produced similar results with anticholinergic drugs in animal experiments since the early years of this century.

WILLIAM MCBRIDE

Foundation 41,
365 Crown Street, Surry Hills,
NSW 2010, Australia

Divine artefact

SIR—If God was smart enough to create The System, he was certainly smart enough to cover his tracks, that is he could have 'implanted' the geological astronomical record so that what many of us now see as a scientifically pre-Creation history is merely a divine artefact.

I say this as a non-creationist scientist who nevertheless cannot find a way around this argument. Can you? If not, then science would also appear to be a religion: we simply believe there was no relatively recent Creation but cannot prove it.

BRUCE DENNESS

Bureau of Applied Science Ltd,
42a High Street,
Newport, Isle of Wight PO30 1SE, UK

OUP's list

SIR—I write to correct several errors in a recent news article (*Nature* 336, 102; 1988). OUP has been a significant publisher of scholarly journals for many years, and before our acquisition of IRL's list of 13 journals we had a list of about 100 journals in the sciences and humanities, such as *Brain* and *Mind*.

We also are successful textbook publishers. We rarely publish theses.

In your issue of 10 November, between pages 118 and 119, the publication of several textbooks is announced, but no theses.

KATHERINE JURY

Oxford University Press,
Walton Street, Oxford OX2 6DP, UK

More sideshows for solar neutrinos

It seems agreed that the numbers of detectable neutrinos emitted by the Sun are smaller than they should be, but people are still vainly searching for an explanation.

THE matter of the missing neutrinos from the Sun becomes ever-more tortuous. It is now well over a decade since R. Davis's experiment, deep underground in the Homestake mine in the United States, first revealed that there are only a third as many neutrinos emitted from the Sun as the supposedly well-understood nuclear physics of the energy-producing core would suggest there should be.

Since then, the chief consequence has been that speculation about the origin of the discrepancy has flourished. But it seems to be generally agreed that further experiments will be needed to tell whether the discrepancy arises because knowledge of the interior of the Sun is inadequate or because neutrinos and their interaction with matter are not fully understood.

Davis's experiment, it will be recalled, consists of a large tank full of tetrachloroethylene (once used for removing grease from people's clothes). Although neutrinos interact only weakly with most particles of matter, they will occasionally convert nuclei of ^{37}Cl into nuclei of ^{37}Ar , which is radioactive and thus detectable in small numbers.

Davis and his colleagues were at the outset confident of their finding that the rate of conversion of ^{37}Cl into ^{37}Ar is roughly a third of what it would be in the light of the expected flux of neutrinos from the Sun. Neither further examination nor the passage of time has seen them waver in this conviction.

Now there has arisen a further source of distraction in a field already sufficiently confused — the possibility that some of the conversion of chlorine to argon nuclei observed originally by Davis may be driven not by neutrinos from the core of the Sun, but by solar flares. The suggestion appears to have been made last year by Davis himself, based on an apparent correlation between records of the Homestake equipment and the presence of flares on the Sun. What seems to have weighed most with Davis is that an experimental run begun during 1972 spanned the occurrence of the great solar flare of August that year and happened to be the run in which chlorine conversion (to argon) was most abundant.

Evidently, if this speculation were correct, the discrepancy between the expected and measured fluxes of neutrinos from the Sun would be further magnified. For the predicted numbers are those expected to be generated by nuclear reac-

tions in the energy-producing core. If flares might be the source of some of the observed neutrinos, the core must be even less prolific a source than had been supposed.

Luckily, but by means of climax followed by anticlimax (*Phys. Rev. Lett.* **61**, 2650 & 2652; 1988) this spectre has now apparently been laid. First, John D. Bahcall from the Institute for Advanced Study at Princeton argues that if the Homestake detector is able to record neutrinos from solar flares, other neutrino detectors now in operation or construction should be even more drastically affected by them.

But how can a mere solar flare perform as if it were material in the high-temperature core of the Sun? One possibility is that nuclear collisions may generate both pions and muons which, when they decay, yield neutrinos detectable at Homestake. The trouble, says Bahcall, is that even the great flare of 1972 would have generated too few of them to be detectable at Homestake in the numbers reported by Davis. On similar grounds, he also rejects the notion that solar flares may influence the Homestake experiment by changing (reducing) the flux of more distant cosmic rays reaching the Earth and thus, indirectly, the amount of artificial radioactivity generated in the high atmosphere, which might indirectly affect the experiment.

But if there is some means by which mesons produced in solar flares can influence the conversion of ^{37}Cl into ^{37}Ar , Bahcall's argument goes, then the other neutrino detectors either built or planned should be even more sensitively affected. Obligingly, he calculates the sensitivity of the two principal detectors now in operation, the Japanese instrument (whose detector is built around 2,000 tonnes of water) called Kamiokande and the Irvine-Michigan-Brookhaven detector (IMB for short), with three times as much water.

Each instrument is designed so that the interaction between a neutrino and a nucleon or electron can be followed as if it were a regular measurement in high-energy physics. Both detectors are now famous for having recorded neutrinos from the supernova event of February, 1987 (11 and 8 events respectively). Bahcall quite rightly says that there should soon be a chance to settle the question whether solar flares can be a source of neutrinos, for is there not a solar maximum on the way?

Solar neutrino physicists, mercifully, will not have to wait even that long. A group of 35 physicists, mostly Japanese but with a strong group from the University of Pennsylvania, report in a succeeding article what amounts to their failure to associate the detection of neutrinos in the Kamiokande tank of water with the occurrence of solar flares. One of the interesting features of the Kamiokande measurements is that it has been possible to use as measures of the intensity of solar flares occurring sporadically not just their visual appearance, or magnetic effects on the surface of the Earth, but the intensity of the proton flux within them as measured by satellites in orbit for this among other purposes.

In all cases, it seems, the measurements at Kamiokande are insignificantly different from the background flux at times of solar flares or even the appearance of γ -ray flux in the galactic cosmic rays. The undramatic but comforting conclusion is that Davis's original discrepancy is not worsened by the suspicion that solar flares somehow play a part in the detection of neutrinos on the surface of the Earth.

Quite what that means for Davis's report last year that there is a correlation remains to be determined. On the face of things, there is no reason why his direct-conversion experiment should be affected by solar flares more than are measurements of the high-energy physics type in Japan and at IMB. But, in a field in which there have already been many surprises, nobody should count on that.

The obvious and long-standing difficulty remains that the solar neutrinos measured by ^{37}Cl conversion span a range of energy different from that to which other detectors would be sensitive (whence Bahcall's scheme for using virtually the whole world's stock of gallium as a detector sensitive in a different energy range). But the possibility also remains that the neutrinos generated in the interior of the Sun are somehow modified in their traverse of the huge envelope around the core, most simply by the conversion of one type of neutrino (the electron-neutrino) into one of the other two (associated with muons and taus, the heavier versions of the electron). However this tale comes out, it will remain a marvel that so much work, experimental as well as theoretical, has been stimulated by a single discrepant observation.

John Maddox

Meteorology

Clouds on the giant planets

Peter J. Gierasch

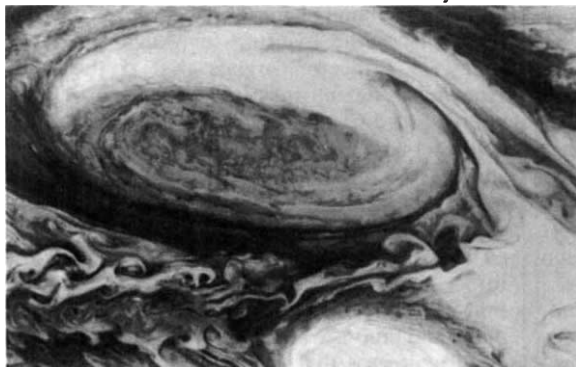
WHAT kinds of clouds exist on Jupiter and the other outer planets? The atmospheres of these planets are principally hydrogen and the trace constituents which can condense to form clouds at observable levels are NH_3 , H_2O , CH_4 and possibly NH_4SH . B. E. Carlson *et al.* ask in a recent paper (*J. atmos. Sci.* **45**, 2066–2081; 1988) whether the various cloud layers are precipitating or not, and if they are, how violent the activity is. They conclude that the visible ammonia clouds on Jupiter (which is too warm for methane to condense; see figure) are not unlike thick cirrus on Earth, but that deeper clouds on Jupiter and all the clouds on Uranus and Neptune could be far more active than any that occur on Earth. These are tough issues, but are extremely important to our understanding of the outer planets. The density and temperature changes that are produced by phase changes and compositional separation within these bodies may have profound influence on the dynamics of their envelopes and on the extrapolation inward of the temperature increase that we observe in outer layers.

Even on Earth where the atmosphere is relatively well documented, the varied nature of clouds is complex and puzzling. At one extreme, there are hazes of very small, slowly growing droplets such as those of photochemically produced sulphuric acid in the stratosphere or in urban smogs. These are nonprecipitating clouds in the sense that coalescence of particles never becomes a runaway process leading to rapid fallout of material. Other clouds are produced by simple condensation of water but are also non- or only weakly precipitating. Radiation fog is an example. Finally, there are actively precipitating systems such as cumulonimbus, in which the heating that arises from the condensing phase change plays an important part in driving fluid motions, so that the existence of precipitation alters the character of the whole system.

To explain this range of behaviour one must know the composition of the atmosphere, including whether grains or ions which can serve as condensation nuclei are present; one must understand the chemistry of the mixture; and one must be able to predict the feedbacks which can occur through radiative or latent-heat effects on fluid motions. It is clear that without observational guidance, discussion of cloud structure and dynamics in remote

and exotic situations is extremely difficult.

Carlson *et al.* wisely restrict themselves to the very simplest considerations. They have evaluated the timescales for the growth of crystals (or droplets), and for fallout and coalescence to occur. No attempt at detailed modelling of the processes is made. The nature of the cloud turns out to depend most importantly on the density of the condensate vapour at the base of the cloud. A low density leads



The Great Red Spot of Jupiter and its neighbourhood, observed during one of the Voyager spacecraft fly-bys. The appearance of the ammonia clouds is consistent with cirrus behaviour, although more detailed observations are needed. The intricate patterns with sharp edges are nicely explained as condensation clouds, forming and evaporating too rapidly for diffusion to destroy the spatial gradients. These clouds occur at a level in Jupiter's atmosphere where the pressure is not far from that at the surface of the Earth and where the temperature is about 150 K. The image is two Earth diameters across and has a resolution of 40 km.

to weakly precipitating clouds because particle size never becomes large enough for differential fall velocities to cause the runaway process of coalescence. A high density leads to precipitation and 'activity'.

Although provocative, this exercise should be viewed with caution. Among the clouds predicted to be heavily precipitating is the Uranus methane system. In this particular case we have data from the Voyager fly-by of 1986 which shows exactly the opposite cloud behaviour. Although the temperature variation with height is nearly adiabatic and thus consistent with strong vertical mixing, a long-lived chemical tracer, the *para*-hydrogen concentration, is vertically layered (R. A. Hanel *et al. Science* **233**, 70–74; 1986), proving that mixing is not occurring. Images also reveal that the planet is covered with uniform haze and not cumulus structures as would be predicted by heavy precipitation and latent-heat release.

Uncertainties in composition are probably not sufficient to be the cause of failure of the predictions. But it is well to bear in mind that among the key condensibles, only methane on Jupiter and Saturn has a well-determined concentra-

tion, and only because it does not condense on those particular planets. The concentration is about 2×10^{-3} on both planets (Gautier, D. *et al. Astrophys. J.* **257**, 901–912; 1982; Encrenaz, T. & Combes, M. *Icarus* **52**, 54–61; 1982), or a bit more than twice what would be expected from a solar ratio of carbon and hydrogen. This and other evidence, such as the bulk densities of the outer planets (Pollack, J.B. A. *Rev. astr. Astrophys.* **22**, 389–424; 1984), are consistent with the relatively rich mixtures of condensibles assumed by Carlson *et al.*, which lead to their prediction of strongly precipitating clouds.

There are a number of factors which could cause clouds to show different behaviour on the outer planets than on

Earth. The condensable matter in the terrestrial atmosphere, H_2O , has a smaller molecular weight (18) than the bulk atmosphere (29) and therefore cannot become stably stratified within a cloud. Recall that the abundance of condensable vapour will always decrease with height in a cloud, because the temperature drops with height and the abundance is limited by the saturation vapour pressure. In contrast, the molecular weights of condensibles on the outer planets are much larger than that of the bulk atmosphere, and stable stratification can develop. Another important difference lies in the absence of a solid lower boundary analogous to the Earth's surface. On Earth, precipitation strikes the surface and cannot be removed any further from the clouds than this. On the outer

planets, precipitation will fall to deeper, warmer levels and evaporate, producing a patch of gas enriched in the precipitating component and heavy relative to its surroundings. Further sinking could ensue, with the end result of removing the condensable gas from the level of mean condensation and reducing its concentration there.

These issues are similar to those encountered in other fluid systems where two constituents are mixed and buoyancy effects arise because of variable concentration. Salinity in the oceans is an example. Such systems have been examined extensively (*Buoyancy Effects in Fluids* by J. S. Turner (Cambridge University Press, 1973) is an excellent introduction). But my feeling is that the prediction from first principles of the behaviour of systems as complicated (and fascinating) as the outer-planetary atmospheres is well beyond our capability and that *in situ* observations are needed. The NASA Galileo probe to Jupiter, scheduled for launch in the autumn of 1989 and to arrive at Jupiter in 1995, may be useful. □

Peter J. Gierasch is in the Astronomy Department, Space Sciences Building, Cornell University, Ithaca, New York 14853, USA.

Palaeontology

Helping with inquiries

Henry Gee and Rory Howlett

THE inquest into what may have been the worst weekend in the history of the Earth continues, with a post-mortem on the events leading up to an incident at the end of the Cretaceous period 65 million years ago, when a force or forces unknown assaulted the biosphere and left much of it in a critical condition. Although some researchers have asked for other, similar events to be taken into consideration (such as the extinction event at the end of the Permian period, 248 million years ago), they are reluctant to press charges. Despite the identity parade lined up for inspection at a recent meeting*, it is hard to pin the blame for particular mass extinctions on individual suspects, be they meteoritic, environmental or biotic. But several have been taken in for questioning and are helping palaeontologists with their inquiries.

Some victims are not always as innocent as they seem. In what might be termed contributory negligence, some species leave themselves wide open to attack; a theme explored by David Jablonski (University of Chicago). Genera whose constituent species have wide geographical distributions, or which contain many different species, are less likely to be picked off by snipers than those confined to small areas or whose specific representatives are few. John Maynard Smith (University of Sussex) looks at species rather than genera, which also sow the seeds of their own destruction. Parthenogenetic species, for example, sacrifice evolutionary longevity for short-term gains in abundance, whereas haplodiploid groups such as social insects last long enough to mother whole dynasties. Such species-selective characters are important contributors to the choice presented to selection in any given situation. But sometimes the assault is so heavy that no species, no matter how adaptable, can avoid it. Jablonski's end-Cretaceous mollusc genera suffer extinction to an equal degree, irrespective of their speciosity, geographical spread or other extinction-proof insurance policies.

Even so, the ecology of a group will make it more sensitive to some factors than others when faced with extinction. It is therefore no surprise that the fortunes of marine plankton, for which fossil data are extraordinarily good, tend to mirror general oceanographic changes more than anything else (Andrew Knoll, Harvard University). The end-Cretaceous event made its presence felt here too,

just as with other marine and terrestrial groups. Plankton are special, though, having been arguably the first group of organisms ever to have been mugged: they underwent at least one main extinction and subsequent adaptive radiation in the Precambrian (earlier than 600 million years ago). But the keynote of these and subsequent extinctions in plankton is that many such extinctions are environmentally induced. The realization of the importance of the environment in extinction could do for palaeobiology what plate tectonics did for geology 20 years ago (Knoll).

Ammonites are also in tune with their environment, and their diversity rises and

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Ammonites dominated the seas for 300 million years, only to become casualties of the end-Cretaceous extinction event 65 million years ago. Only their abundant fossils, such as this example of *Dactyloceras commune* from the Jurassic, betray the existence of this once-successful group.

falls with sea level (Anthony Hallam, University of Birmingham; Michael House, University of Hull), although oceanic anoxia may have seen off a few on occasion. But it is hard to explain why no new ammonoid forms appeared in the 18 million years before the end of the Cretaceous, unless prophecy was an attribute of this diverse and long-lived group.

That environmental factors predetermine the complexion of mass extinctions also leaves little room for doubt in the mind of Richard Fortey (British Museum, Natural History), who has been piecing together the events surrounding two particular mass-extinction episodes. The event at the Cambrian–Ordovician boundary (about 500 million years ago) was a 'watershed' in the evolution of many marine forms, with extinction accompanied by the evolution of new lineages to replace the old. Many of the misfortunes befell epicontinental groups sensitive to sea-level change. These were replaced by

new forms derived from deep-water groups that moved to the continental shelves. But the reverse happened at the Ordovician–Silurian boundary (440 million years ago), when deep-water taxa suffered preferentially. The high rates of extinction in trilobites with pelagic larvae point to a decrease in oxygen tension being responsible for many losses. But the extinctions among trilobites, graptolites, conodonts, brachiopods and other groups were not simultaneous, and may have been related to various phases of growth and shrinkage of the contemporary South Polar ice cap, centred on West Africa.

Terrestrial organisms such as tetrapods are harder to fit into the general scheme than marine groups, perhaps because their record as fossils is so much poorer. Michael Benton (The Queen's University, Belfast) has reassessed A. S. Romer's 20-year-old tetrapod phylogeny data, using modern cladistic techniques and improved relative and absolute dating, in an attempt better to define what is known about the terrestrial response to mass extinctions. Although the best way to advance knowledge is to find more fossils, the most important improvements in our understanding of tetrapod phylogeny in the past 20 years have come from the reorganization of known data by objective taxonomic techniques. This comes despite assertions that paraphyletic groups, being essentially expressions of organizational grade, reflect environmental change better than monophyletic groups. But there are still enormous gaps in the record: for example, the Aalenian Stage (188–181 million years ago) in the Jurassic has left no trace of tetrapods anywhere in the world. This does not mean that all tetrapods became extinct 188 million years ago, only to re-evolve spontaneously 7 million years later. They were presumably there all the time, but the fossil record is so capricious that none has yet been found.

Of all tetrapods, dinosaurs have been linked most closely with mass-extinction events. But the lack of a suitable definition of the Cretaceous–Tertiary boundary applicable to the whole world makes it difficult to assess the time at which the last dinosaurs disappeared (Alan Charig, British Museum, Natural History). The assumed synonymy between the eclipse of the dinosaurs and mass death more generally has clouded the issue; many other animal groups are believed to have suffered hardly a jolt from the presumed bolide meteor impact. And the possibility that some dinosaurs survived the event cannot yet be excluded. The case of the dinosaurs is an example of a more general problem, that of establishing synchronicity of events on a geological scale (C. H. Holland, Trinity College, Dublin); for a mass extinction to be meaningful, its duration needs to be specified together

Jane Burton/Bruce Coleman

*Joint Royal Society and Linnean Society discussion meeting *Evolution and Extinction*, London, 9–10 November 1988.

with the litany of its victims.

Vascular plants provide good forensic evidence from the Cretaceous of western North America. Local floras show that perhaps 46 per cent of palynomorphs disappear at the Cretaceous-Tertiary boundary, ferns becoming common immediately afterwards in some sections. Leaf and cuticle fossils show a gradual change from moderately dry conditions in the pre-impact world to much wetter ones afterwards (R. A. Spicer, Goldsmith's College, London). But floral change need not be connected with impact and, as with dinosaurs, the evidence can only be applied locally or regionally rather than globally. For example, the 75 per cent drop in leaf macrofossil diversity at the boundary layer in New Mexico is matched by a 50 per cent drop in Montana, and a fall of less than 40 per cent in Alberta, whereas the event hardly registers in the Southern Hemisphere. This evidence corroborates geological indications that the meteorite landed somewhere in the North American area. But this could be because North America is where the problem has been studied most intensively, and here at least the evidence points to some ecological catastrophe, leading to sudden vegetational changes across the boundary.

Mammals, too, show that when mass extinctions are studied in detail, the results say more about prevailing local conditions than the broad, planet-wide march of events. The mass extinction in the middle of the Miocene Epoch (24.5–5.1 million years ago), recognized in both the tetrapod and planktonic foraminiferan records, is to some extent reflected among the rodents of the Iberian peninsula studied by Jean-Jacques Jaeger (University of Paris VI), but most of the changes in diversity reflect local immigration and evolutionary responses to isolation on islands, and are of purely regional significance. We know from studies of natural extinctions on islands (M. H. Williamson, University of York) that species turnover — extinction and immigration — is common, and that permanent extinctions are rare. So although a regional perspective on faunal and floral change shows up more detail than a global view, little is learned about mass extinctions from the adoption of too parochial an attitude.

Palaeoenvironmental markers, then, may be more revealing than faunal or floral turnover to those seeking to prise changes in global conditions from the rocks. Changes in the isotope composition of fossils and the sediments surrounding them are coming to the fore in this respect. Deep ocean temperatures are related to changes in the relative abundances of isotopes of oxygen. Not only do these changes illustrate periodicity over the short term, such as the progress of the Ice Ages over the past 2 million years, but they also illustrate periodicities of much

lower frequency. Cycles between 20 and 40 million years may reflect climatic changes associated with continental drift (N. J. Shackleton, University of Cambridge). Although palaeotemperature on its own is unlikely to affect the longevity of more than one or two foraminiferan species at a time, the synchronously vanishing foraminifera at the Cretaceous-Tertiary boundary speak of an event that is unusual "by any standards", unlikely to be a statistical artefact. Isotope composition of marine sediments may also reflect other changes important to the biosphere, such as the changing oxygen tension of sea water. This is just one of several factors that may cause extinctions independently (Antoni Hoffman, Palaeobiology Institute, Warsaw). The chance concatenation of several could lead to a mass extinction.

David Raup (University of Chicago), an architect of the periodic mass-extinction hypothesis, discussed attempts to come to terms with stratigraphical incompleteness, using the ammonite

record from the Cretaceous-Tertiary boundary of northern Spain as an object lesson in how to determine whether extinctions were sudden or gradual. As so tellingly shown with the tetrapod record, incompleteness distorts our perception of events. False trails and missing clues can make sudden extinctions appear gradual, or vice versa. The problem is hard to resolve, but it seems that diversity loss can trace its roots to a combination of factors that have both sudden and gradual effects.

Although the evidence to incriminate a bolide is mounting in the case of the Cretaceous mass extinction, there is no evidence for extraterrestrial impacts in any other mass extinction, with one or two possible exceptions. And Jared Diamond (University of California, Los Angeles) pointed out that most of the observed species extinctions within the past few centuries have been caused, either directly or indirectly, by the effects of man. □

Henry Gee and Rory Howlett are on the editorial staff of Nature.

Helioseismology

Deep roots of solar cycles

Douglas Gough

MOUNTING evidence indicates that the entire Sun partakes in the solar cycle, the 11-year periodic variation in solar activity. Thus it became clear at a recent meeting* that the 11-year changes in sunspot activity are related to large-scale variations that perhaps penetrate to the very core of the Sun where the thermonuclear reactions take place, and are not confined to surface and sub-surface effects. This forces solar physicists to consider the possibility that the solar cycle is essentially a global oscillation and is not entirely the product of a dynamo operating in the turbulent convection zone immediately beneath the surface of the Sun.

The solar cycle last peaked in 1981, reached a minimum in 1986 and is now increasing again. Activity is typically revealed by the number of sunspots present at any time but large-scale acoustic oscillations, the subject of the meeting, are another indicator. Just as geoseismology elucidates the inner structure of the Earth, so helioseismology reveals the inner structure of the Sun.

Thus, S.M. Jefferies and co-workers (National Solar Observatory, Tucson) have recently remeasured two coefficients that describe the acoustical properties of the Sun which depend on its departure from spherical symmetry. The frequencies of the acoustic waves can be expressed as an expansion in Legendre polynomials, whose argument is $m/\sqrt{l(l+1)}$ where l

is the degree of the mode (the number of nodal lines on the surface) and m is the azimuthal order (the number of nodes around a line of latitude). The lowest term depends on the purely spherical aspects of the solar structure and the higher order terms result from corrections accounting for the departure from spherical symmetry.

The first coefficient measured by Jefferies and co-workers, a_2 , describes the lowest-order even correction (the quadrupole component). This has been declining since the last solar maximum in 1981, but the latest measurement, made in 1987, which was reported by Jefferies, is the first since the solar minimum, and shows an upturn. Interestingly, the next even coefficient, a_4 ,

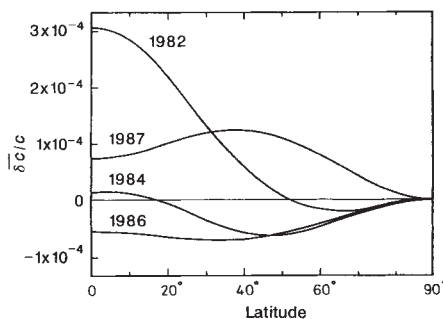


Fig. 1 Asphericity of the Sun as measured by the difference in speed of sound $\delta c/c$ of modes which would be 'degenerate' (have identical frequencies) if the Sun were spherical. The curves were derived from the coefficients of the Legendre polynomials (discussed in the text) describing the frequency splitting.

* Seismology of the Sun and Sun-like stars. Tenerife, 26–30 September 1988.

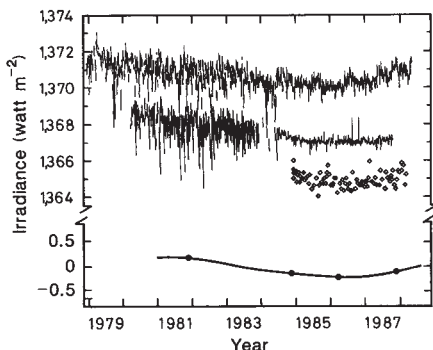


Fig. 2 Daily mean solar irradiance measured by the earth radiation budget experiment (ERBE) on Nimbus 7 (top curve), the Active Cavity Radiometer on the Solar Maximum Mission (second curve) and the ERBE on the Earth Radiation Budget Satellite (diamonds). The smooth curve at the bottom (note the change of scale) is computed from the sound-speed variations depicted in Fig. 1, assuming that they result solely from variations in temperature, but accounts for only a third of the observed change. The other two thirds (5×10^{-4}) presumably result from a luminosity variation.

has also been declining since 1981, but shows no sign yet of reversing its trend. Thus not only the magnitude, but the shape of the acoustical distortion is changing (Fig. 1). In 1987, the distortion was maximum in the mid-latitudes where sunspots of new polarity first appear in the cycle.

How deep-seated is the asphericity? In principle, the variation with depth can be ascertained from how the coefficients a_l vary with l . This is because the depth to which the modes sample increases as l declines¹. But the data are hard to obtain from ground-based measurements, and the scatter is large. It seems that the asphericity declines slowly with depth, but the data could be consistent with the asphericity being constant down to 0.45 solar radii, which is the base of the cavity confining the deepest modes ($l=25$) for which the a_l are available at a time of high solar activity. Evidently one needs to study modes of yet lower degree to learn about the inner core. This is a challenging task, partly because the number of modes available to measure the asphericity increases almost linearly with l , so at low l there is less scope for averaging the data in order to reduce the noise.

What is the origin of the acoustical asphericity and how is it related to the magnetic cycle? Although the non-uniform solar magnetic field inhibits small-scale convection in the surface layers of the Sun, and thereby changes the thermal stratification, that change is too small to explain the observed splitting in the oscillation frequencies. But the change could be amplified by a resultant large-scale circulation (a similar effect occurs in uneven cooling ponds). Parker² has proposed that a circulation is driven by magnetic fields concentrated at the base of the convection zone which might alternatively account for the asphericity (J. Kuhn, Michigan State University). In either case, magnetic field and angular momentum would be advected by the flow, leading possibly to coherent latitudinal variations of angular velocity and magnetic field in the photosphere. A very impressive correlation presented by E. Ribes (Observatoire de Meudon) lends observational support to this idea, which may also relate to the so-

called torsional oscillations observed by LaBonte and Howard³.

If the subphotospheric acoustical asphericity is naively interpreted as a latitudinal variation in the speed of sound on spherical surfaces, and is smoothly extrapolated to the surface, a pole-equator difference in effective temperature is inferred. Its magnitude, as Kuhn has pointed out⁴, is broadly consistent with direct observation. Moreover, by redistributing the radiative flux from the photosphere, the temporal variation modulates the 'solar constant' (the solar irradiance as detected at Earth), even

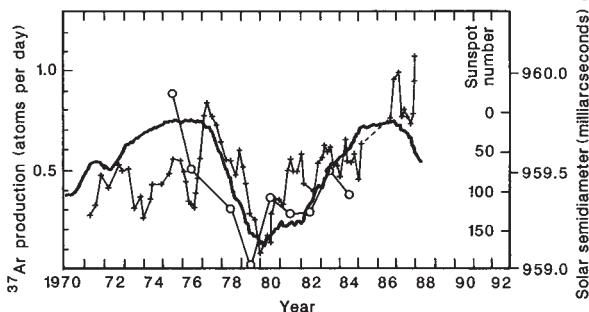


Fig. 3 Correlations of sunspot number (solid curve), solar neutrino flux recorded by R. Davis Jr and co-workers (crosses) and apparent semidiameter of the Sun, from astrolabe measurements (circles), perhaps indicating a causal connection.

at constant luminosity (Fig. 2). Although the calculated effect has the observed form, the amplitude is much too low. Does this mean that the latitudinal variation in effective temperature is really much greater than that inferred from the speed of sound, conflicting with Kuhn's analysis⁴, or is there a temporal variation of the solar luminosity too? That would imply a deep-seated origin of the solar cycle.

Parker's ideas² of magnetic shadowing at the base of the convection zone would lead to a temporal luminosity variation. So would a modulation of the core. If the correlations between sunspot numbers, apparent diameter and neutrino flux (Fig. 3) are symptomatic of a causal connection, then it is almost inevitable that the nuclear reaction rates in the core are varying with the cycle (E.A. Gavryuseva and T. Zaitsepin, Institute for Nuclear Research, Moscow). It is unlikely that the neutrino modulation could result from changes in the solar envelope. Thus it seems plausible that there are dynamical variations in the heart of the Sun. Oscillatory aspherical motion in the core and its

environs have been indicated by the interesting profile of angular velocity⁵ inferred from rotational splitting of 5-minute oscillations. Recent discussions of whether the long-term phase coherence of the cycle is too high to be compatible with a turbulent dynamo, and is instead suggestive of a regular oscillation of the radiative interior^{6,7}, relate to this idea. Moreover there now seem to be discrepancies in the speed of sound, density and hydrogen abundance between standard solar models and those values inferred from the many seismic observations reported. These discrepancies may well be a consequence of cyclically varying motion, but they can be partially accounted for (S. Turck-Chieze, CEN-Saclay) in the new standard solar model, constructed with a superior equation of state, that is discussed by Christensen-Dalsgaard and colleagues on page 634 of this issue⁸.

How should the low-degree oscillations that penetrate to the energy-generating core be affected? Their frequencies would be reduced by 0.8 microhertz (μHz) if the large-scale horizontal thermal stratification at the base of the convection zone were changed sufficiently to account for a decline of 5×10^{-4} in relative luminosity (Fig. 2). A corresponding change to the core would yield a decline of 0.1 μHz , whereas an adjustment to the surface layers would have little effect. There were several reports at a meeting (*Advances in helio- and astroseismology*, Aarhus, 7–11 July 1986) of a decline of around 0.4 μHz since the solar maximum, but one of an increase.

Some of the new measurements reported at the recent meeting confirm the 0.4- μHz decline but others do not, leaving the issue unresolved. This highlights how difficult it is to interpret the observations. Often the frequency splitting by rotation and acoustical asphericity is unresolved, or contaminated by noise and the products of inadequate observing windows, which lead to mode beating and amplitude changes that shift the apparent mean frequency of the modes contributing to a peak in a power spectrum. However, some analyses of the low-degree data show that the frequency variation depends on the value of l . This fits the asphericity variation shown in Fig. 1, but only if this penetrates deep into the Sun. □

1. Christensen-Dalsgaard, J. *et al.* *Nature* **315**, 378–382 (1985).
2. Parker, E.N. in *The Internal Solar Angular Velocity* (eds Urney, B.R. & Sofia, S.) 289–301 (Reidel, Dordrecht, 1987).
3. LaBonte, B. & Howard, R.F. *Solar Phys.* **80**, 373 (1982).
4. Kuhn, J.R. *Astrophys. J.* **331**, L131–L134 (1988).
5. Duvall, T.L. Jr *et al.* *Nature* **310**, 22–25 (1984).
6. Dicke, R.H. in *Stellar Rotation* (ed. Slettebak, A.) 289–317 (Reidel, Dordrecht, 1970).
7. Bracewell, R.N. *Q. J.R. astr. Soc.* **29**, 119–128 (1988).
8. Christensen-Dalsgaard, J., Däppen, W. & Lebreton, Y. *Nature* **336**, 634–638 (1988).

Douglas Gough is in the Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA, UK.

Molecular genetics

Fruitful fruitflies in Crete

Michael Ashburner and Peter Lawrence

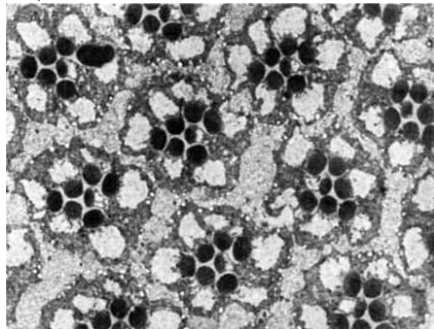
SINCE 1978, a revolution has occurred in *Drosophila* biology, a revolution that stems from an enormously fruitful interaction between classical genetics and molecular biology. Only ten years ago we were arguing about restriction maps of the heat-shock protein genes; cloning a gene is now the least of our problems — indeed we heard during a recent meeting* that a consortium has begun to construct a complete overlapping map of the *Drosophila* genome with cosmids. Soon, we hope, cloning by phoning will be a reality for all drosophilists.

The molecular characterization of genes is but a prelude to a diverse approach to a molecular understanding of biological processes. That said, however, some very remarkable things are being discovered with respect to gene structure in *Drosophila*. Perhaps the most remarkable was that reported by R. Davis (University of Texas, Houston) for the gene *dunce*, which encodes a cAMP phosphodiesterase. This gene was first identified because flies carrying mutant alleles are stupid, and they flunk certain associative learning tests. The protein-coding region of this gene is not unusual, yet the gene encodes at least ten overlapping RNAs (there are at least 13 exons) that cover more than 100 kilobases. Most remarkably, one of the introns includes not one but at least two, and probably more, unrelated genes, including the well-known *Sgs4* gene which encodes a salivary gland glue protein. What can possibly be the functional significance of such a complex organization of a housekeeping gene? One answer, of course, is none. Comparative studies of this gene in related species may help to provide a clue, as will, in due course, transformation of the cAMP phosphodiesterase coding region with various combinations of 5' exons.

The genetic and molecular analysis of early development in *Drosophila* is at an exciting stage. The anterior-posterior axis of the embryo can be perturbed by mutations in several different genes. One, *bicoid*, encodes a nuclear protein distributed in an anterior-posterior gradient at blastoderm formation (C. Nüsslein-Volhard, Tübingen); another, *vasa*, which affects the posterior pole rather than the anterior, encodes a protein involved in determination of germ cells (Y. Jan, University of California, San Francisco).

The patterns of cell division in the post-blastoderm *Drosophila* embryo are highly ordered. In a careful study, V. Foe

(University of Washington) has shown that the embryo is divided into a set of bilaterally symmetrical domains, within each of which mitosis is nearly synchronous. Foe suggests that division of the blastoderm into these mitotic domains indicates the commitment of cells to specific developmental fates. One protein, the product of the gene *string*, is required for the entry of nuclei into division. During the syncytial cleavage divisions this is achieved by using maternally inherited *string* messenger RNA. But the onset of post-blastoderm mitosis requires the zygote to become independent of its mother, and make its own



Cross-section of the eye of *Drosophila* showing the arrangement of seven photoreceptor cells (cell number 8 is below the plane of the section). Cell number 7, missing in the *sevenless* and *boss* mutants, is that which is located centrally.

string product. Then, the pattern of cells that express the *string* RNA is very similar (if not identical) to the pattern of mitotic domains seen by Foe. Interestingly, the *string* gene encodes a protein which resembles that encoded by *cdc25* of fission yeast, a gene acting as a dose-dependent activator of mitosis (P. O'Farrell, University of California, San Francisco).

Drosophila biologists have a new weapon in their armoury — the constructs of C. O'Kane and W. Gehring (Biozentrum, Basel) that express the *Escherichia coli* β -galactosidase gene under the control of a weak promoter. The expression of this enzyme, conveniently stained blue or detected with an antibody, is very sensitive to the position within the genome at which a vector carrying it inserts. Using the innovative methods for generating new transpositions of P-element insertions developed by A. Spradling (University of Baltimore) and W. Engels (University of Wisconsin, Madison), a very large number of strains, each with this construct in a different genomic site, are now being analysed. Some (a surprisingly high proportion) show an 'interesting' pattern of β -galactosidase expression. For example, one studied by C. Goodman (University of California, Berkeley) has a pattern of expression that

resembles that of the cell-adhesion molecule fasciclin III. Indeed, this insertion is just 5' to the gene encoding this protein.

The homoeodomain proteins are one of the more dramatic discoveries of the drosophilists. Scott and Gehring suggested some years ago that the homoeodomain looks like a DNA-binding protein of the class exemplified by the lambda repressor, and this now appears to be so. NMR studies by Gehring and colleagues show that the homoeodomain does form a helix-turn-helix broadly similar to (but in detail subtly different from) the lambda repressor. On the subject of homoeoproteins, we note that D. Hogness (Stanford University) has now shown, using specific antibodies, that the *Ubx* gene encodes a family of proteins that are germ-layer specific in their expression. One, for example, is expressed in relatively late embryos (and in larvae), but only in the central nervous system.

The development of the ommatidia of the eye (Benzer's neurocrystalline lattice) is looking more interesting by the minute. Tomlinson and Ready had already suggested that the eight photoreceptor cells of each ommatidium are recruited in a fixed order surrounding cell eight (see figure). At the meeting, E. Hafen (University of Zurich) discussed his studies of the *sevenless* gene, known to be required for the development of cell seven (but no other). For the correct development of cell number seven, *sevenless* must be active in this cell itself. The *sevenless* gene encodes a cell-surface protein that is a tyrosine kinase: a single amino-acid substitution within the ATP-binding site of the kinase domain inactivates the *sevenless* gene. If this gene product is a receptor, what is its ligand? A possible answer is the product of a gene, *bride of sevenless*, whose mutant phenotype is very similar to that of *sevenless*. L. Zipursky (University of California, Los Angeles) has now shown, by mitotic recombination, that *boss* is required not in cell seven, the cell missing in mutant flies, but in cell number eight, a cell not visibly affected by the *boss* mutation. Perhaps *boss* encodes a ligand on the surface of cell eight that induces the neighbouring cell seven to differentiate.

A similar picture of cell-cell interaction in the developing eye is emerging from a study of a gene called *rough*. The product of this gene is required only by cells 2 and 5, yet in a *rough* mutant these cells are recruited normally into the pattern — it is only cells 3,4,1,6 and 7 (which do not require *rough*) that fail to differentiate. The *rough* protein includes a homoeobox and is, as would then be expected, a nuclear protein (G. Rubin, University of California, Berkeley). The picture emerging is one of specific signals emanating from the first cells that differentiate, which then trigger the determination of neighbouring cells. The identification of the

*Fifth EMBO *Drosophila* workshop, organized by S. Artavanis-Tsakonas, Kolymbari, Crete 28 August – 4 September 1988.

100 years ago

EDISON'S PERFECTED PHONOGRAPH



THE marvellous results attained by Mr. Edison's recent improvement on the original phonograph of 1878 have induced us to present a view of the latest form of the instrument.

The illustration shows the reproducing diaphragm in position, the operator listening through the tubes, and standing behind the instrument so as to be seen.

From *Nature* 39, 107; 29 November 1888.

molecular signals will be a step to understanding these processes at the cellular level.

The hormonal control of gene activity has long fascinated developmental biologists, especially when the consequences of hormone action are as dramatic as they are in the higher insects — controlling the metamorphosis of a maggot (an expressive word, usually banned from polite drosophilists' conversation) into a winged fly. A few genes are known to respond very rapidly to ecdysteroids, steroid-moulting hormones, and many years ago it was suggested that some of these genes encode proteins that not only turn themselves off but also turn on later-responding genes. This may indeed be so, at least as a first approximation, for Hogness and C. Thummel (University of Utah) have now characterized the products of two of the so-called early ecdysone-responsive genes. Both are DNA-binding proteins and one (the product of the *E75* gene) encodes proteins that have both zinc fingers and domains that resemble those of vertebrate steroid and retinoic-acid receptors. R. Hill (CSIRO, Sydney) has shown that antibodies against the *E75* protein associate with both the *E74* and the *E75* early genes in polytene chromosomes, as well as at a developmentally changing set of sites which include those of some late genes. But there are still surprises in store for the study of the ecdysone response — for example, P. Cherbas (Indiana University) has found that one of the early ecdysone-inducible genes of tissue-culture cells (*eip40*) encodes a γ -cystathionase. □

Michael Ashburner is at the Department of Genetics, University of Cambridge, CB2 3EH; Peter Lawrence is at the MRC Laboratory for Molecular Biology, Cambridge CB2 2QH, UK.

Spectroscopy

New aspects of Raman scattering

Robin M. Hochstrasser

RAMAN spectroscopy, relying on the inelastic scattering of light, is a widely accepted analytical tool. In particular, it is regularly used to identify molecular species, their structures or internal energy distributions, where other techniques fail. Indeed, at a recent conference* devoted to the method, 475 papers were presented, most concerning new studies using the Raman effect as a tool. Some, however, described new aspects of the physics of the effect, particularly time-dependent ones resulting from molecular-dynamic interactions and which might be revealed using femtosecond lasers. This illustrates how, 60 years after its discovery, and 100 years after its discoverer's birth (marked in *Nature* with extracts¹ from his acoustical studies), the Raman effect continues to stimulate fundamental physical research.

The essential aspects of Raman scattering experiments are well understood (see figure). A laser at frequency ω_L is focused into the sample while the frequency distribution of the scattered field (ω_S) is examined by means of a monochromator. Emission lines appear at energies $\hbar(\omega_L - \omega_i)$ corresponding to molecular transitions whose polarizabilities can oscillate significantly at these frequencies. The dynamical characteristics of the sample are contained in the intensities, shapes and polarization properties of the Raman transitions.

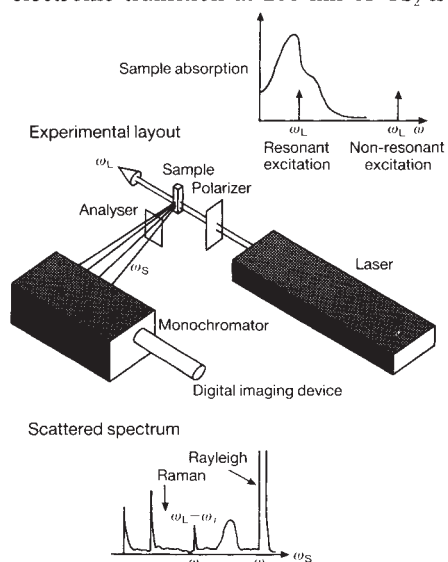
The width is a measure of the dephasing rate of the molecular vibrations: the rate of random interruptions in the coherent evolution of the field at ω_S that result from the scattering molecule interacting with its surroundings. Attempts to elucidate liquid and solid-state dynamics require a fuller understanding of such ultrafast molecular relaxation processes. It is now clear that Raman spectroscopy can contribute substantially to this, particularly in combination with advances² with femtosecond lasers.

The bound states of atoms in collision-free conditions undergo spontaneous emission as the only mechanism irreversibly to extract energy from a radiation field. This process is Raman or Rayleigh ($\omega_S = \omega_L$) scattering. Collisions between atoms allow energy to be exchanged with the field so that the spontaneous emission following near-resonance excitation can become shifted from the laser frequency to produce, in time, a fluorescence or collisionally redistributed emission. For molecules, the collisions also generate a population distribution of the many nearby excited states,

leading to a third type of spontaneous emission.

In condensed matter, where all these processes occur, resonant Raman scattering is easily identified because of the energy-conserving relationship ($\omega_S = \omega_L - \omega_i$) of its narrow line emission and the laser frequency. The response time of Raman scattering is determined by the difference in the periods of the light and the material electronic oscillations or by the dephasing rate of the optical transition that is being driven if the excitation is resonant. If the resonant Raman scattering from bulk matter were time-resolved, the signal would decay with the same effective dephasing time that determines the growth of the fluorescence signal. These times are probably in the range of 10–100 femtoseconds (fs) for liquids and could be studied by pulsed-laser methods.

The absolute cross-section (probability) of resonant light scattering also depends on the rate at which the resonant optical transition is dephased or relaxed. A knowledge of the radiative emission rate, which depends on the integrated absorption, and the yield of Rayleigh plus Raman scattering thus provides a measure of the rate of dephasing in the optical transition. Also, the scattering efficiency can be used to determine whether the optical absorption band is inhomogeneously broadened over the timescale determined by the dephasing. These concepts were recently brought out and applied quantitatively to the CS_2 molecule by Myers and co-workers³ who show that in hexane solution at 300 K the electronic transition at 200 nm of CS_2 is



Standard layout of a Raman-scattering experiment, with typical absorption and scattered spectra as inserts (see text).

* Eleventh International Conference on Raman Spectroscopy, London, 5–9 September 1988.

characterized by a dephasing time of 65 fs and that the solution consists of a homogeneous distribution of scatterers.

Ultrafast processes can also be exposed by studies of the Raman polarization (using various polarizer/analyser combinations; see figure). A traditional concept is that the scattering initiates from a fixed, random distribution of molecules because the Raman lifetime is shorter than vibrational or rotational periods. The usual polarization rules are therefore not applicable to resonant phenomena where the Raman lifetimes are expected to be relaxation dependent. The degradation of polarized signals from rapidly rotating molecules has long been used to clock rates of photodissociation⁴ and similar clocking experiments were recently done using resonance Raman excitation profiles of gases⁵. We have shown⁶ that even in the rapid-relaxation environment of a liquid, the polarization anisotropy of resonance Raman signals may be used to specify dephasing rates. The resonance Raman anisotropy calculated for an aniline molecule shows that the vibrational memory and the average molecular rotational period at 300 K are long enough for significant depolarization to occur before the Raman polarizability oscillation is dephased by collisions in the liquid.

Molecular dynamics is also concerned with the flow of population from one state to another. Such processes influence Raman scattering and can be studied directly using time-resolved Raman scattering or infrared spectroscopy. Results show that molecular vibrational energy can be stored in liquids or solids often for many hundreds of picoseconds⁷⁻¹⁰. Chemical processes should be influenced by these energy bottlenecks which exist for much longer than a local liquid structure.

The Raman transitions from many liquids and solids are homogeneously broadened, that is to say the local structures in the material are exchanged sufficiently frequently for all molecules to contribute to the Raman signal at any frequency, ω_s . But in our recent time-resolved Raman-scattering study¹¹ of pyrrole in a glass and in solution we found

that the population relaxation of the N-H stretch vibration does depend on which portion of the vibrational transition is initially excited. This suggests that models of relaxation should incorporate a distribution of structures, each with its own relaxation parameters — a more complex view of liquid dynamics than usually presented. The structures in this example are believed to have an N-H bond directed into the π -cloud of a nearby molecule. Evidence for this is drawn from the remarkable propensity for energy transfer between N-H and ring vibrations that is found in this system — about 60 per cent of all the energy in N-H vibrations flows into ring stretches¹¹.

Time-resolved Raman methods are also used to follow the course and identify intermediates in fast chemical reactions.

The ultimate time resolution is limited by the Raman dephasing time that clocks the relevant part of the spontaneous emission in such experiments. Although nanosecond and picosecond transient Raman effects are now frequently used¹² in molecular studies, only recently have these methods been extended into the femtosecond regime¹³. Contrary to usual expectations, the spectral resolution of such experiments is not in principle limited to the Fourier-transform width of the laser pulse (a few hundred wave numbers for a 50-fs pulse). New types of Raman spectroscopy are being developed in which the spectral widths are entirely determined by the kinetics of the system. □

Robin M. Hochstrasser is in the Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

Linguistics

Genes and the tower of Babel

Jared M. Diamond

WILL it ever be possible to deduce the full branching tree of historical relationships among the 5,000 or so extant languages in the world? At first, this goal seems utterly hopeless. Human language may have existed for several hundred thousand years, yet methods for recognizing language relationships may fail after only about 10 millennia of language divergence. There is not even an agreed classification of extant languages. Recent assaults on these problems by linguists and archaeologists can now draw on the findings of geneticists as well. Since 1963, Cavalli-Sforza and colleagues have been analysing modern gene distributions to reconstruct the phylogeny of extant human populations. Their latest effort¹ examines frequencies of 120 alleles in 42 aboriginal human populations.

Cavalli-Sforza *et al.* could recognize six main population clusters, among which the oldest split is between Africans and the five non-African clusters (compare with conclusions based on mitochondrial DNA², discussed in News and Views³ last year). The next split divided a pair of clusters consisting of Caucasoids and north-east Asians/Amerindians from a trio of clusters consisting of south-east Asians, Pacific islanders and New Guineans/Australians. There were subsequent splits within the pair and trio and within each of the six clusters. Archaeological data tentatively suggest dates of 15–90 kyr (15,000–90,000 years ago) for the early splits among the six main clusters.

For each of the 17 language phyla recognized by Ruhlen⁴, it is striking that aboriginal speakers of all the languages in that phylum mainly belong to the same genetic cluster recognized by Cavalli-Sforza and

colleagues. Initially this seems surprising when one thinks of the many modern cases in which languages have been transferred with little transfer of genes. For instance, the people of Papua New Guinea have negligible European admixture but have adopted English as their national language. But most such transfers probably depended on a degree of political organization that emerged in the past 5,000 years. Before then, the match between linguistic phyla and genetic clusters would have been undisturbed.

Evaluation of these proposed correlations between genes and languages requires the solution of three sets of issues. One obvious issue at the root of any proposed comparison of linguistic and genetic classifications is the validity of the linguistic classification used. Even within certain regions of the world, notably New Guinea and the Americas, there seemed until recently to be many languages and stocks with no known relation to each other, let alone to the stocks elsewhere in the world. Proposals for order in this chaos have been made by Wurm⁵ for Papuan languages, and by Greenberg⁶, who sees all Amerindian languages as falling into only three phyla. Ruhlen and Greenberg classify world languages into 17 and 15 main phyla, respectively.

These recent classifications of languages require further testing. For example most of the approximately 740 identified Papuan languages were formerly believed to be unrelated to each other, but recent research has permitted them to be grouped into about 11 phyla or macro-phyla and 8 isolated languages. Nevertheless, it still involves an act of faith to assume further that these 11 Papuan phyla

1. Bowsher, J.M. *Nature* **336**, 17, 109, 206 & 310 (1988).
2. Shank, C.V. in *Generation of Ultrashort Optical Pulses* (ed. Kaiser, W.) 5–34 (Springer, Berlin, 1988).
3. Myers, A.B., Li, B. & Ci, X. *J. chem. Phys.* **89**, 1876–1886 (1988).
4. Yang, S. & Bersohn, R. *J. chem. Phys.* **61**, 4400–4421 (1974).
5. Ziegler, L.D. *J. chem. Phys.* **86**, 1703–1714 (1987).
6. Myers, A.B. & Hochstrasser, R.M. *J. chem. Phys.* **87**, 2116–2121 (1987).
7. Laubereau, A. & Kaiser, W. *Rev. mod. Phys.* **50**, 607–665 (1978).
8. Seilmeier, A. & Kaiser, W. *Rev. mod. Phys.* **50**, 279–317 (1978).
9. Velsko, S. & Hochstrasser, R.M. *J. phys. Chem.* **89**, 2240–2253 (1985).
10. Heilweil, E.J., Cavanagh, R.R. & Stephenson, J.C. *J. chem. Phys.* **89**, 230–239 (1988).
11. Ambroseo, J.R. & Hochstrasser, R.M. *J. chem. Phys.* **89**, 5956–5957 (1988).
12. Atkinson, G.H. (ed.) *Time Resolved Vibrational Spectroscopy* (Gordon & Breach, London, 1987).
13. Petrich, J.W. *et al. Biochemistry* **26**, 7914–7923 (1987).

and 8 Papuan isolates all belong to a single world-level phylum of Ruhlen's or Greenberg's classification.

A second issue is to develop improved methods for quantifying relations among languages, and for detecting relations between distant languages⁷. Linguists face the task of classifying entities that evolve at different rates, are described by hundreds of characteristics (vocabulary and grammar), and could come to resemble each other through independent convergent evolution or borrowing as well as by shared ancestry. Much progress has been made in solving the precisely analogous tasks that face geneticists like Cavalli-Sforza and colleagues, and that also face comparative anatomists, community ecologists and students of protein evolution. Linguists have not reached consensus on the most appropriate methods and might gain from the quantitative statistical methods used in these other fields.

The remaining issue concerns the time depth of human language. If language in substantially its present form has existed for several hundred thousand years, it would be hopeless to trace all extant languages to a common ancestor. But language as we know it may go back only to the origin of anatomically modern *Homo sapiens* (200–130 kyr?) or even of behaviourally modern *H. sapiens*⁷ (50 kyr?). As Cavalli-Sforza *et al.* note¹, the advantage conferred by complex modern language may itself have been important in the rapid expansion of behaviourally modern *H. sapiens* over the world, and in the approximately simultaneous disappearance of Neanderthals, who may have had cruder languages.

There is, of course, a pessimistic alternative: languages may have great time depth, expansions of the first farmers may have eliminated most earlier language phyla, and the surviving phyla may be unrelated. But if the time depth of modern language is not so great, and if improved methods for discerning relations between languages that diverged more than 10,000 years ago can be developed, the problem of tracing relations among surviving languages may be soluble. To quote Greenberg⁶, "The ultimate goal [of historical linguistics] is a comprehensive classification of what is very likely a single language family. The implications of such a classification for the origin and history of our species would, of course, be very great". □

1. Cavalli-Sforza, L.L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 6002–6006 (1988).

2. Cann, R.L. *et al.* *Nature* **325**, 39 (1987).

3. Wainscoat, J. *Nature* **325**, 13 (1987).

4. Ruhlen, M. *A Guide to the World's Languages* Vol. 1 (Stanford University Press, 1987).

5. Wurm, S.A. *Papuan Languages of Oceania* (Narr, Tübingen, 1982).

6. Greenberg, J.H. *Language in the Americas* (Stanford University Press, 1987).

7. Klein, R.G. *A. Rev. Anthropol.* **12**, 25–48 (1983).

Jared M. Diamond is a professor of Physiology at the University of California Medical School, Los Angeles, California 90024–1751, USA.

Molecular evolution

rDNA world falling to pieces

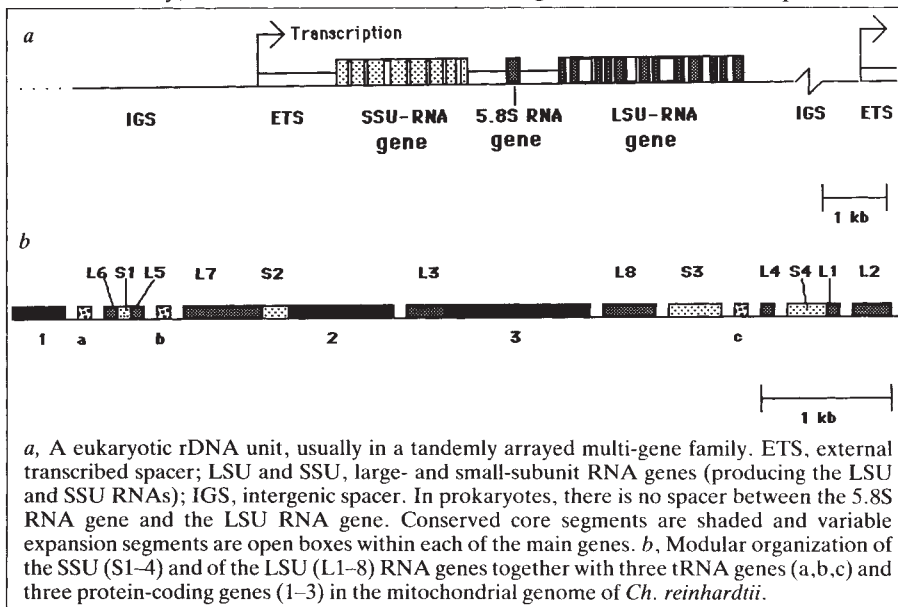
Gabriel A. Dover

If physics has its laws, biology has its variety. Which is to be expected given the messy march of evolution. Nothing would seem more predisposed to generate a mess — a sort of molecular roulette in action — than the recent findings^{1,2} by P. H. Boer and M. W. Gray of a bizarre fragmentation and scrambling of scraps of ribosomal RNA genes (rDNA) among protein-coding and transfer RNA genes. Yet nothing of the sort happens; instead, perfectly assembled complex two-dimensional (and by inference three-dimensional) RNA structures, identical to the conserved rRNA 'core' structures within the large and small ribosomal subunits, emerge by intermolecular interactions between the unspliced mini-gene transcripts.

Until recently, it has been taken almost

events in the 'RNA world' and the intermediate steps in the evolution of genes and introns are becoming clearer.

A comparison of rRNA sequences among species has shown that they consist of modules of conserved core segments interspersed with more variable segments (also called 'expansion segments')^{5,6}. Until recently, in all species that have been studied the core and variable segments are co-transcribed to form the main RNA species of the large and small ribosomal subunits. This picture changed, however, with a study of the nuclear rDNA and rRNA of the trypanosoid protozoan *Crithidia fasciculata*⁷. At the DNA level in this species, there is the usual arrangement of sequences encoding conserved segments of large-subunit RNA interspersed with



for granted that if any part of the genetic system were to resist the variety of events imposed on protein-coding genes: fragmentation into exons and introns; somatic rearrangements; differential transcription; and differential *cis* and *trans* splicing, it would be the rDNA whose products are at the heart of the ubiquitous translation machinery and whose genomic organization into a repeating unit consisting of a small number of genes and spacers is generally invariant among the main kingdoms (*a* in the figure). But the discovery of catalytic RNA and the suggestion that rRNA itself has a catalytic role⁸, together with the suggestion¹ that the first ribosome consisted solely of RNA, has brought the study of rDNA and rRNA into the limelight (for review see ref. 5). This has led to the analysis of these systems in unfashionable organisms, from which the early

more variable sequences. These are co-transcribed as usual, but unlike other species that have been studied, many of the variable segments are removed from the primary large-subunit rRNA transcript of *Cr. fasciculata* to generate covalently unlinked fragments. The fragments representing the core sequences and a few remaining variable sequences can then, in principle at least, assemble and fold into the usual rRNA pattern of secondary and higher-order structure.

Are the modules the relics of ancient mini-genes (DNA or RNA) whose descendants are now physically linked, co-transcribed and folded *ensemble* together with the remains of their intervening spacer sequences, as economical solutions to the problems of living in a modern, stressful world? Support for this suggestion comes from new work on the

mitochondrial genome of the ciliate protozoan *Tetrahymena pyriformis*⁸ and the unicellular green alga *Chlamydomonas reinhardtii*^{1,2} in which even more bizarre permutations have been uncovered. In *T. pyriformis* there are two large-subunit rDNA modules (separated by a tRNA gene) whose order is the reverse of that expected from their order in the RNA; and in *Ch. reinhardtii* there are four small-subunit and at least eight large-subunit rDNA modules interspersed with each other and with various protein-coding and tRNA genes (*b* in the figure). These are also 'illogically' scrambled among themselves. As before, abundant and separate large- and small-subunit mini-RNAs emerge from the single long transcript of the complex DNA unit, after removal of flanking sequences to form two networks typical of small- and large-subunit RNA secondary structures.

Boer and Gray^{1,2} point to an additional clue to unravelling rDNA evolution which derives from the potential of one of the three protein-coding genes (*b* in the figure) to give rise to a product that could function either as a reverse transcriptase or as an RNA maturase, given the similarity in sequence with open reading frames in group II introns involved with RNA processing in fungal mitochondria. The first activity suggests that the scrambled arrangement is a more recently derived feature reflecting haphazard insertions of complementary DNAs; the second could have been part of the machinery for the evolutionary removal of extraneous RNA lying between the modules, if indeed modules were once independently existing mini-genes. The first entertains the possibility that complex rearrangements were spread by internal (molecularly driven) events, whereas the second might have been a streamlining response to external (selective) pressures, extreme examples of which are the rDNAs in mammalian mitochondria which have lost several 'expansion segments'.

The concept of early modular organization predicts that first, each DNA module might have had a separate promoter and terminator; and second, enticingly, that each RNA module might have existed within separately functioning ribonucleoprotein particles not unlike present-day small nuclear RNAs and their associated proteins. With regard to the first prediction, in some organisms

the chief rRNA genes (*a* in the figure), including the 5S rRNA genes, can be separately transcribed, leading to the proposal⁹ that these were once independent units. These might have been intermediate steps in the process of assembling the smaller modules into more manageable units, as indeed might be the longer L7 and L8 pieces (*b*) consisting of mixtures of core and variable segments which are not further processed at the RNA level. The evolutionary end-point of this process would be present-day long rRNA genes co-transcribed from a single promoter whose variable (expansion) segments are unprocessed relics of intervening spacers that have become an integral part of large- and small-subunit rRNA secondary structures. With regard to the second prediction, a relationship between small eukaryotic and prokaryotic RNAs has been suggested recently by the finding¹⁰ of primary sequence and secondary structural similarities between 7S RNA and *Escherichia coli* 4.5S RNA.

As with all good biological stories, the situation is more complex than this. Experimental manipulations of expansion segments in yeast by R. Planta (personal communication) show that they are not all dispensable in the assembly and functions of ribosomes, although they lie in peripheral regions of the large- and small-subunit RNA secondary structures and do not seem to be associated with proteins. Furthermore, despite wide fluctuations in size and composition, often due to the continual gain and loss of short repetitive motifs of DNA by slippage, they form remarkably conserved secondary structures within kingdoms (as a consequence of compensatory mutations occurring in key stem regions)^{11,12} and are also coevolving as a set within any one species⁶. Their evolution might not be explicable solely on the basis of their being molecular fossils of unremoved rDNA spacers. There are arguments⁵ that some of them are more recent acquisitions, like many introns in protein-coding genes. Whatever the case, it is clear that there are some rDNA evolutionary paradoxes still looking for a solution. □

1. Boer, P.H. & Gray, M.W. *Cell* **55**, 399–411 (1988).
2. Boer, P.H. & Gray, M.W. *EMBO J.* **7**, 3501–3508 (1988).
3. Cech, T.R. & Bass, B.L. *A. Rev. Biochem.* **55**, 599–629 (1986).
4. Crick, F.H.C. *J. molec. Biol.* **38**, 367–379 (1968).
5. Gerbi, S.A. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **LII**, 709–719 (1987).
6. Hancock, J.M. & Dover, G.A. *Molec. biol. Evol.* **5**, 377–391 (1988).
7. Spencer, D.F., Collings, J.C., Schnare, M.N. & Gray, M.W. *EMBO J.* **6**, 1063–1071 (1987).
8. Heinonen, T.Y.K., Schnare, M.N., Young, P.G. & Gray, M.W. *J. biol. Chem.* **262**, 2879–2887 (1987).
9. Clark, C.G. *J. molec. Evol.* **25**, 343–350 (1987).
10. Poritz, M.A., Strub, K. & Walter, P. *Cell* **55**, 4–6 (1988).
11. Michot, B. & Bachelier, P.P. *Biochimie* **69**, 11–23 (1987).
12. Hancock, J.M., Tautz, D. & Dover, G.A. *Molec. biol. Evol.* **5**, 393–414 (1988).

Gabriel A. Dover is in the Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK.

Daedalus

Organic-vapour farming

PLANTS, says Daedalus, can extract a lot more from the atmosphere than carbon dioxide. Some roadside plants are usefully fertilized by the nitrogen oxides they capture from vehicle exhausts; and one report claims that office plants like chrysanthemums and spider-plants combat 'sick building syndrome' by taking up solvent vapours, formaldehyde, and carbon monoxide from the underventilated air. In principle, any enterprising green plant should gladly abandon photosynthesis to exploit such a ready-reduced source of carbon. So DREADCO's biologists are passing dilute radiolabelled organic vapours like alcohol and methane over a variety of plants, from tomatoes to rhubarb to privet, looking not just for absorption but for metabolic uptake. Now most plants, like most animals, have a resident flora of bacteria; in at least one case (the nitrogen-fixing bacteria in legumes) this relationship has deepened into a useful symbiosis. So to help things along the DREADCO team are dusting their plants with the methanotrophic bacteria which metabolize methane, simple hydrocarbons, and some of their derivatives. These organisms may have unsuspected collaborative talents. Some have recently been found living symbiotically in the gills of mussels near offshore oil platforms, helping the mussels to digest the local hydrocarbons. With good fortune, they will do the same for Daedalus's plants.

Daedalus has high hopes of his organic-vapour-metabolizing plants. Not only will they deodorize the depressing effluvia of high-tech offices, the fug-laden air of crowded parties, and the ominous smell of dentists' waiting rooms; they will clean up industrial air. As a final purification filter, a greenhouse of the new plants will wonderfully purify the slightly smelly and contaminated air discharged by most industrial processes. The resulting output of country-fresh air and useful vegetables will be a double improvement.

Even better, Daedalus hopes to develop strains of organic-trapping grass, shrubs, hedge plants, and so on, for planting in the environment at large. Particularly in polluted regions, they should have a powerful advantage over less accomplished species, and should spread vigorously. In bright sunlight they will photosynthesize carbon dioxide; under smoky skies they will make do with the smoke. At last the grim effluvia of industry will be integrated into nature. Daedalus even hopes that his plants will help to counter the growing contribution of atmospheric methane to the greenhouse effect. But even he has little hope of developing a strain of plant which will metabolize fluorocarbons, and save the ozone layer.

David Jones

Corrigendum

In the News and Views article by M.F. Perutz (*Nature* **336**, 202; 1988) it was implied that in a dimer of identical subunits related by a 2-fold symmetry axis, equivalence of the subunits is preserved only by rotations about and shifts along axes that are perpendicular to the 2-fold axis. In fact this is true for any pair of axes (one axis for each subunit) that are related by the same 2-fold symmetry axis as the subunits themselves.

Earthquakes and urban growth

SIR—The disaster in Armenia is a reminder that earthquakes are most damaging when their stored energy is released near cities, even when those cities are of no great size. Earthquakes cannot be prevented and cities cannot be relocated. Hence, the buildings in cities that may be so afflicted must be appropriately designed. In developed nations, where memories of great urban earthquakes are vivid (San Francisco, 1906; Tokyo, 1923), collaboration between seismologists and urban planners has led to the implementation of measures to mitigate the effects of future earthquakes. In developing nations such collaboration may not exist because of uncontrolled urban development, the lack of informed awareness about earthquake risk or economic circumstances that mean earthquake risk has a low priority.

Earthquake risk is especially high in the

developing nations, because of poor construction methods, and will be considerably higher in the future. The process of urbanization coupled with rapidly increasing populations will generate more than a hundred cities with populations exceeding two million by the year 2000 (ref. 1). By a geographical coincidence 40% of these cities are to be found within 200 km of a tectonic plate boundary or close to an earthquake that has caused damage in the past (Figs 1 and 2). It appears that within 12 years 290 million 'supercity' dwellers, 80% of them in developing nations, will live in a region of seismic risk (Table 1). Because the mean growth rate of large cities in the developing nations is currently more than 2% per year, we can expect the populations at risk to double before 2035.

Damaging earthquakes at some active

plate boundaries may recur in less than 50 years, at others after more than 200 years. Typically we cannot forecast earthquakes more accurately than within a few decades². We make a distinction between long-term forecasts and short-term predictions³. Short-term predictions of earthquakes cannot prevent the collapse of buildings; a long-term forecast, on the other hand, provides time to introduce appropriate measures to mitigate its effects⁴.

At its most basic the earthquake cycle is a process of the build-up of strain and its release. Hence measurements of displacement, strain and tilt provide fundamental data for monitoring progressive deformation in the region of a future earthquake. Space technology provides us with powerful new techniques for monitoring the surface deformation of the Earth in three dimensions. Very-long-baseline interferometry and satellite laser ranging have confirmed the instantaneous velocities of the world's tectonic plates, hitherto inferred from geological data, while Global Positioning System (GPS) geo-

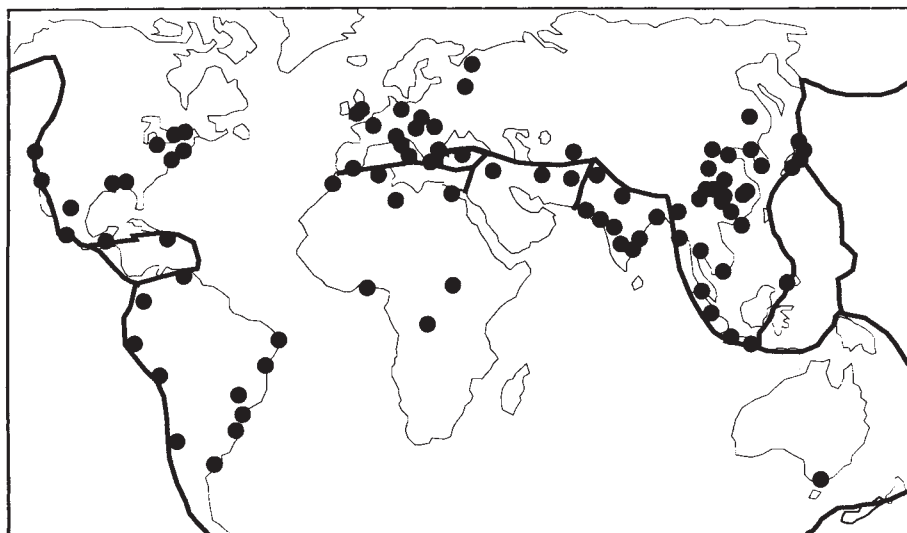


Fig. 1 Distribution of cities with a projected population of two million people or more in the year 2000. Bold lines represent convergent or strike-slip boundaries.

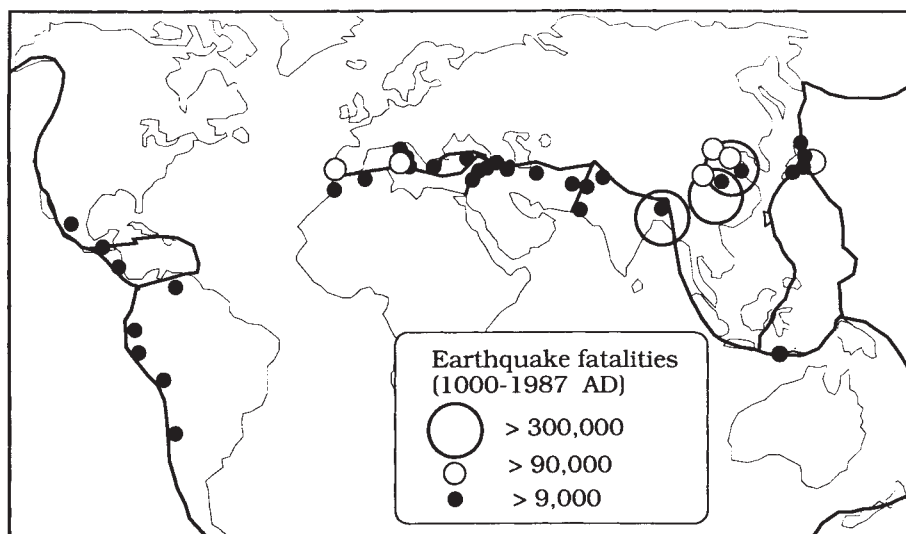


Fig. 2 Distribution of earthquakes in the past 1,000 years in which over 9,000 people died. Most such disasters have occurred where cities coincide with plate boundaries.

Table 1 Population growth of cities of population two million or more by the year 2000 that are located within 200 km of a potential magnitude 7 earthquake or a historically damaging earthquake

City	Year 1975	Year 2000*	Growth rate (%)
Mexico City	11.61	25.8	2.2
Tokyo, Japan	16.37	20.2	0.5
Tehran, Iran	4.39	13.6	3.4
Jakarta, Indonesia	5.53	13.2	3.2
Delhi, India	4.63	13.2	3.5
Karachi, Pakistan	4.03	12.0	3.7
Dacca, Bangladesh	2.34	11.2	5.2
Beijing, China	8.91	11.2	1.8
Manila, Philippines	5.04	11.1	2.8
Los Angeles, USA	8.96	11.0	0.3
Bangkok, Thailand	4.05	10.7	3.6
Osaka, Japan	7.96	10.5	0.6
Tianjin, China	7.43	9.7	1.9
Lima, Peru	3.7	9.1	2.8
Baghdad, Iraq	2.72	7.4	3.1
Bogotá, Colombia	3.02	6.5	1.7
Lahore, Pakistan	2.42	6.2	3.4
Medan, Indonesia	0.91	5.4	5.2
Shenyang, China	3.53	5.3	2.2
Santiago, Chile	3.38	5.3	1.3
Ankara, Turkey	1.67	5.2	3.3
Caracas, Venezuela	2.61	5.0	1.7
Algiers, Algeria	1.58	5.1	3.7
Rangoon, Burma	1.76	4.3	3.2
Naples, Italy	3.8	4.3	0.4
Guadalajara, Mexico	1.95	4.1	2.3
Chengdu, China	2.0	3.9	2.5
Taipei, China	1.84	3.7	2.6
Surabaya, Indonesia	1.81	3.7	3.0
San Francisco, USA	3.1	3.5	0.4
Chongqing, China	2.59	3.3	2.0
Istanbul, Turkey	2.87	3.3	1.2
Kanpur, India	1.51	3.2	3.0
Athens, Greece	2.25	3.0	0.8
Xian, China	1.95	3.0	2.3
Singapore	2.26	2.9	0.8
St Domingo, Dominican Rep.	1.14	2.9	2.8
Kabul, Afghanistan	0.7	2.8	4.3
Kitakyushu, Japan	1.74	2.2	0.4
Guatemala City	0.85	2.1	3.7
Nagoya, Japan	1.94	2.0	0.0

*Populations (in millions) approximately double in the last quarter of this century and will do so again in developing nations in the following 35 years¹.

desy offers similar data faster and less expensively³. The system is similar in accuracy to the best available terrestrial geodesy, but far surpasses it in ease of application and production rate. Yet in the developing nations geodesy remains poor or non-existent. The longer precise geodesy is delayed in such nations, the longer will be the delay in developing earthquake forecasts based on epicentral strain development. One step forward was the recent adoption by an ICSU sub-committee of a resolution that space geodesy should be applied to the characterization of seismic risk in the developing world.

Populations in some earthquake-prone regions will more than double within less than the time of one complete earthquake cycle. Space geodesy offers a method for long-term forecasting that will identify areas that are especially at risk, and in

which stringent building regulations, especially for high-rise constructions, should be imposed and observed. No 'supercity' was affected in the Armenian disaster yet the death toll was enormous because of inadequate building practices. A combination of new technology and recognition of the centres of population especially at risk can help reduce the scale of such catastrophes in the future.

ROGER BILHAM

*CIRES, University of Colorado,
Boulder, Colorado 80309,
USA*

1. *Prospects of World Urbanization Revised 1984-5* (UN Dept Int. Econ. & Soc. Affairs, 1987).
2. Nishenko, S.P. & McCann, W.R. *Maurice Ewing Series* Vol. 4, 20-28 (American Geophysical Union, 1981).
3. Rikitake, T. *Earthquake Prediction* (Elsevier, 1976).
4. Ambraseys, N.N. in *Seismicity and Seismic Risk of the Offshore North Sea Area* (eds Ritsema, A.R. & Gurrin, A.) 317-345 (Reidel).
5. Bilham, R. in *Aerospace Century XXI 1637-1659* (American Astronautical Society, 1988).

Trial of human malaria vaccine

SIR—I wish to raise an important point in connection with the report by Patarroyo *et al.*¹ of the first human trial of a vaccine, which is composed of synthetic peptides, against the asexual blood stage of the malaria parasite, *Plasmodium falciparum*.

Patarroyo *et al.* evaluated humoral immune responses in their volunteers before immunization and challenge. For that they used both an ELISA test, taking an optical density of above 0.2 as positive, and an IIFA test, which was taken to be positive at titres above 1:20. By those criteria, 7 of the 13 volunteers were positive by one or both tests before immunization. In what way may the antibodies detected in these 7 volunteers have influenced the immune response to vaccination?

More important, why were the antibody-positive volunteers not excluded from the trial on the grounds that the presence of antibodies indicates past exposure to malaria? Malarial infection is very common in the Colombian military forces. Many blood donors at the Central Military Hospital are asymptomatic malaria carriers² who are not aware of their past infection.

Those involved in malaria vaccine development should be very careful to design protocols adequate for evaluation, and take great care not to make over-optimistic claims based on preliminary results. Even if a vaccine becomes available, it would be some years before it could be integrated into the multiple approaches needed to control this complex disease.

CARLOS A. ESPINAL

*Sociedad Colombiana de Parasitología y
Medicina Tropical, AA 100084,
Unicentro Bogota, Colombia*

SIR—Several points made by Patarroyo *et al.*¹ deserve more careful discussion.

First, the immunization protocols differed for the two different proteins. Ought the different protection results be attributed to the differences in proteins, vaccina-

tion schedules or times before challenge? The authors do not comment on the fact that two doses of Spf(60)30 gave better protection than three.

Second, cross-reactivity in the humoral response to the two very different protein immunogens was detected in several vaccinees. If this is explained by a reaction against the linking peptide, as the authors propose, it would follow that there was almost no immune response against important epitopes in most vaccinees.

Third, the authors mention that very different results were obtained depending on whether acridine orange or Giemsa were used to determine parasitaemia. Which were reported?

MOISES WASSERMAN

*National Institute of Health,
Avenida Eldorado con Carrera 50,
AA 80080, Bogota, Colombia*

PATARROYO REPLIES—In response to the question of whether some of our volunteers had previously had malaria, we have now tested both by ELISA and IIFA the sera from 109 volunteers, all from non-endemic areas (which was one of our selection criteria) and who had never been in malarious areas. As with the volunteers in our vaccine study, about 10 per cent of the new sera had very low (less than 1:20) antibody titres by IIFA. This is accepted throughout the world as background level in this type of test. Malaria infection is common within the Colombian military forces but not within members coming from non-endemic areas.

What is important is our demonstration of an increase in antibody titres upon immunization although, as we and others have found in the past few months, there is no correlation between antibody titres and protection. It is worth adding that in all cases our original volunteers had a negative peripheral blood mononuclear

cell stimulation index when tested with schizont sonicates before immunization, further indicating that the volunteers had never been in contact with malaria.

With regard to the question of different vaccination schedules for the two proteins, I cannot agree that it is possible to distinguish whether two doses are better than three from the data we present.

The cross-reactivity mentioned by Wasserman could be accounted for by the NANP sequence in both peptides rather than the linker peptide; in any case, as already mentioned, protection does not seem to be correlated with humoral and cellular responses.

Finally, as stated in our paper, we used the much more sensitive acridine orange method rather than Giemsa to test parasitaemia. During the challenge we employed the test every 12 hours, and every 4 hours on critical days, so that the volunteers would never be at excessive risk.

MANUEL ELKIN PATARROYO

*Immunology Institute,
Hospital San Juan de Dios,
Bogota, Colombia*

1. Patarroyo, M. E. *et al.* *Nature* **332**, 158-161 (1988).
2. Espinal, C. A. & Morales, P. O. *Trans. R. Soc. trop. Med. Hyg.* **78**, 645-647 (1984).

Whale size quandary

SIR—One cannot easily accept Downhower and Blumer's penultimate statement¹ that "if an aquatic neonate is smaller than this lower limit, then it cannot generate metabolic heat fast enough to compensate for heat loss" without qualification. One of the reasons that a large whale has such a low basic metabolic rate is that the circulation time of its blood volume is many minutes, making the delivery rate of nutrients to its cells a slow process^{2,3}. The smaller neonatal whale will have a much shorter blood volume circulation time, a higher basic metabolic rate, and a cardiovascular system that could be physically capable of increasing delivery rate to a greater extent than is possible in the adult⁴.

Provided nutrient availability to the neonate itself does not become limiting, there seems no reason to suppose that it could not compensate for heat loss by increased metabolic activity. This, however, is unlikely to be achieved without affecting growth rate, which is the other serious drain on resources in early life. Size at birth presumably represents a workable compromise between the 'demands' of the two competing processes of heat production and growth, and the latter cannot be ignored.

DENYS WHEATLEY

*Department of Pathology,
University of Aberdeen,
Aberdeen AB9 2ZD, UK*

1. Downhower, J.F. & Blumer, L.S. *Nature* **335**, 675 (1988).
2. Coulson, R.A., Hernandez, T. & Herbert, J.D. *Comp. Biochem. Physiol.* **56A**, 251 (1977).
3. Coulson, R.A. *Perspect. Biol. Med.* **27**, 121 (1983).
4. Coulson, R.A. *Comp. Biochem. Physiol.* **84A**, 217 (1986).

A passion or a science?

John C. Marshall

Understanding Psychotherapy: The Science Behind the Art. By Michael Franz Basch. *Basic Books: 1988. Pp.329. \$19.95.*

THE claim that psychotherapy is, in part, an art needs no elaborate defence. And in this respect psychotherapy is no different from the practice of any other branch of medicine . . . or teaching, or law or indeed scientific research. In all domains of human endeavour, the exercise of intuition, discretion and empathy constitutes the bedrock of effective praxis.

Michael Franz Basch is a professor of psychiatry in Chicago, and also a training and supervising psychoanalyst, with an extensive private practice. For the most part, his book consists of a series of clinical vignettes of neurotic patients, described in a fashion that is refreshingly free of jargon. These patients appear to have been helped by Professor Basch's artful aid during the course of therapy based on somewhat eclectic principles. Despite his eminence as an analyst, Basch is not a doctrinaire freudian. His book emphasizes "the quest for competence and self-esteem" to a greater extent than did the Master, who was often content to transform neurotic despair into ordinary unhappiness. To judge from these case-histories, Basch is a wise, insightful and humane therapist into whose consulting rooms one would, with some confidence, entrust oneself or one's loved ones, should the need arise. Basch no doubt intends to give precisely such an impression, but this is no cause for scepticism; books by psychiatrists that provoke the contrary reaction are not hard to find.

Most of the problems that Basch illustrates are commonplace. The everyday round of people not coping as well as they might with the pressures of work and family; the vicious circle of anxiety; the inability to give and receive affection; the feeling of not being wanted and appreciated; sex and religion (too little, too much or not the right sort)—these are the despairs that fill the waiting room. And in Basch's presence they seem to yield to a commonsensical 'talking cure'. But the sense of obviousness is, of course, available only in retrospect and was doubtless not accessible to the patient's own unaided efforts. It is always easier to have superb insight into someone else's problems; one's own foibles and manifest inadequacies tend to be veiled to introspection: hence psychotherapy.

So far so good. Basch has produced an impressive, albeit tacit, justification of the value of psychotherapy as an art. My

puzzlement and unease arise from the book's subtitle, *The Science Behind the Art*. In his introduction, Basch writes: "The building blocks for the construction of a valid, unifying, and useful explanatory theory for psychotherapy lie all around us — one need only put them in place, and then demonstrate how such a theory may be profitably applied". These are ambitious, not to say foolhardy words in the current state of our knowledge. They positively invite Karl Kraus's riposte that "Psychoanalysis is a passion, not a

"I fear that the prestige of science has outstripped its achievements in relation to our understanding of human beings. There *are* branches of psychology that truly merit the designation 'scientific', but they are few and far between."

science; it lacks the steady hand of the investigator".

One might reasonably expect that a psychiatrist who wanted to refute Kraus's aphorism would begin by demonstrating that psychotherapy actually worked. It is minimally incumbent upon Basch to provide evidence, of a moderately objective character, that patients who receive professional psychotherapy do become better (or better faster) than those who do not. No such evidence is given; indeed the very question is never raised. The notion of 'spontaneous recovery' appears to have no place in Basch's vocabulary. If it were true that professional help helped, one would next need to know whether such therapy was demonstrably more effective than the benefits to be derived from intensive discussion with, for example, an old friend, one's priest or rabbi, or a random stranger met on a long train journey. Again, Basch provides no evidence or even discussion of the point. But if the answer were once more positive, one would like to know why. Does psychotherapy work (if it does) because the therapist is especially experienced in dealing with human enigmas, because of the aura attached to his or her professional role, because he or she is the only person who has the time to devote an hour or so a week to listening to one's complaints, because one is paying good money to have someone listen. . . ?

Again, such questions are conspicuous by their absence. Finally, one might

consider the issue of whether psychotherapy is useful because the therapist has privileged access to a well-motivated theory of the human psyche. Is that putative theory grounded in scientific discovery and does it go well beyond what your or my grandmother has intuited over the course of a life that has seen its fair share of the slings and arrows of outrageous fortune? On this question, Basch does have something to say:

I replace Freud's idea — that the underlying motivation for all behavior is the discharge of *instinct* — with the striving for *order, competence, and self-esteem*, a predisposition that is already experimentally demonstrable in infancy; while Freud suggested that a *psychic energy* is responsible for the intensity of thought, I attribute that intensity to the force of information in the form of *affect, feeling, emotion, or empathy*; where Freud spoke of *psychic structure*, I talk about *feedback and pattern matching*; and instead of falling back on Freud's postulation of a *mental apparatus* composed of imaginary agencies like *ego and id*, I refer to the *decision-making functions of the brain and to the self-system*. . .

Oh dear, oh dear. This is exactly what my grandmother would have said before she read Freud. The 'science' Basch serves up is good, old-fashioned folk psychology, supplemented with a little engineering terminology. 'Feedback' has a fairly precise meaning when used of the governor of a steam turbine, but in Basch's writing it is just a fancy synonym for 'Keep your eyes and ears open'. Basch's science consists of some nice pictures of an infant smiling in response to its mother's smile, a spiral drawing labelled "self-esteem" at the bottom and "competence" at the top, and a simple flow-chart in which the boxes have such inscriptions as "pattern of expectation" and "decision making". It's difficult to know what all this is, but science it isn't. Freud did at least provide an alternative mythology in quasi-scientific garb, a set of alternative legends to the conventional fables of religion and philosophy. Basch wants the glory of science, but none of the pain of rigorous hypothesis construction and testing. In this, he is perhaps more freudian than he knows.

Why should an intelligent psychotherapist feel compelled to dress up his perfectly respectable practice in such ridiculous remarks as: "The self system, initially a creation of the brain, comes, through its mediating function, to control and guide the brain in its relation to the world"? Surely an honest therapist is worth his or her weight in gold and does not need to parade the trappings of brain science? Unless, of course, it is the brain that is to be treated. Some of Basch's patients were taking a variety of sedatives and tranquilizers, drugs that Basch himself prescribes sparingly and only after

due thought. But again, his book provides little account of the relationships between psychopharmacological and purely psychological intervention. Lithium *per se* may not 'cure' an affective disorder, but there must, one presumes, be guidelines for how it could be used effectively in association with counselling? Some of these issues are broached in the sections on depression, but the level of discussion is low even when the outcome seems sensible enough: "My rule of thumb is that any patient sick enough to require psychotropic medication deserves concomitant psychotherapy and follow-up".

The whole concept of neuropsychiatry receives disappointingly short shrift in a book that aspires to bridge "a long-standing and counterproductive gap between psychology and biology". And nowhere is the shrift shorter than in the paragraphs on electroconvulsive therapy, one of the unfortunate side-effects of which is memory loss. Here Basch writes: "My experience, admittedly not recent,

has been that the permanent memory loss in patients receiving electroconvulsive treatment was only for the details of the illness and the treatment itself". This claim may, or may not, be true, but Basch should surely have provided a brief review of the experimental evidence on the topic.

I fear that the prestige of science has outstripped its achievements in relation to our understanding of human beings. There are branches of psychology that truly merit the designation 'scientific', but they are few and far between. Basch's book provides no grounds to include psychotherapy among them. Which is a pity, if for no other reason than that good psychotherapy from a practitioner as competent as Basch could in fact be extremely valuable. But the case might have looked stronger (and the book have been a hundred pages shorter) without the pseudo-science. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford OX2 6HE, UK.

Cladistic futures

Philip D. Gingerich

The Phylogeny and Classification of the Tetrapods. Vol. 1 Amphibians, Reptiles, Birds; Vol. 2 Mammals. Edited by Michael J. Benton. Oxford University Press: 1988. Vol. 1 pp. 377, £50, \$98; Vol. 2 pp. 329, £45, \$90.

REVOLUTIONS in science are like waves crashing on a shore. They arise quietly, gain energy and eventually break thunderously, spraying showers of new perspective that alter our view of the shoreline forever. For the past 20 years, a new wave, cladistics, has been building in systematic biology. It is finally breaking, producing phylogenies and classifications in profusion, and after all the rhetoric we can finally begin to evaluate the results.

The Phylogeny and Classification of the Tetrapods represents part of the proceedings of an international symposium convened in London in 1987 and sponsored by the Systematics Association, the Linnean Society and the Palaeontological Association. Each volume opens with a brief preface by the editor, stating that the intention of the symposium was to try to present balanced coverage of the systematics of all the main tetrapod groups. Reading between the lines, it seems that balance was soon lost in controversy over methods and interpretative details. There are 20 chapters altogether (nine in Vol. 1, eleven in Vol. 2) written by 34 contributors, virtually all of whom are palaeontologists. Cladistics is now established around the world and readers are mercifully spared the all-too-familiar primer on

the inadequacy of the fossil record, parsimony, sister-groups, three-taxon statements and so forth. There are no real surprises, apart from Bishop and Friday's α - and β -haemoglobin sequences suggesting that birds and mammals should be grouped together.

How has cladistics changed our view of the history of life? *The Phylogeny and Classification of the Tetrapods* gives us a chance to compare tetrapod phylogeny today with our understanding before cladistics became widely applied. The main differences from pre-cladistic Romer (as seen in his book, *Vertebrate Paleontology*, published by the University of Chicago Press in 1966), result from a new and sometimes questionable willingness to make precise claims about sequences of branching that occurred over short intervals of geological time (often at or within gaps in the fossil record).

Prizes for data matrices go to Gauthier, Kluge and Rowe (14 amniote taxa $\times$ 94 characters, analysed using PAUP), Evans (54 diapsid taxa $\times$ 226 characters, analysed by hand), Novacek, Wyss and McKenna (partial matrix of 20 living mammalian orders $\times$ 67 characters, analysed in pieces by hand, PAUP or PHYSIS), Andrews (10 primate taxa $\times$ 54 characters, analysed using PAUP), Prothero, Manning and Fischer (13 ungulate taxa $\times$ 28 characters, explored using MacClade), and Gentry and Hooker (39 artiodactyl taxa $\times$ 116 characters, analysed using PAUP).

Consistency indices of cladograms based on full matrices range from Gauthier *et al.*'s 0.71 (29 per cent of state changes incongruent with proposed cladogram), to Andrews's 0.64 (36 per cent incongruence), to Gentry and Hooker's 0.29

(71 per cent incongruence). Bishop and Friday caution that comparing and evaluating all possible tree diagrams is not possible because of limitations on computing time, Gauthier *et al.* analyse 14 taxa because this is all their mainframe PAUP can branch and bound, and Novacek *et al.* use software limiting them to full analysis of trees for a maximum of nine taxa. The remaining authors can have little idea whether theirs is the most parsimonious tree.

These volumes reflect the state of cladistics in the late 1980s. The good news is that authors of four out of the 20 chapters (20 per cent) use molecular or morphological information in conjunction with a rigorous algorithm to produce cladograms from data. The first of the bad news is that authors of the remaining chapters (80 per cent) produce cladograms by hand in unknown ways, using characters just to diagnose the branches. The next bad news is that consistency indices peak at 71 per cent (29 per cent incongruence) — character convergence, parallelism and reversal are common in evolution. But the worst news is that maximum parsimony solutions cannot be computed for problems involving more than 14 or so taxa (these problems are 'NP-complete' in the language of computer science, which means difficulties of combinatorial optimization make them computationally intractable). The index to the first volume lists 800 vertebrate taxa, that to the second volume 800 more. No one will ever compute a maximum parsimony solution showing what we really want to know — the genealogical relationships of these 1,600 taxa (or more). And if we can't do it with a computer then we can't do it by hand.

The problem is that we expect too much of morphology in asking it to tell us the genealogy of organisms as well as what they look like. We ask too much of form that it should tell us time. Fortunately there is an independent record of time in the geological/stratigraphical record. Romer took time from geology rather than morphology, and it looks like he got the main outline of tetrapod evolution right. If each author here and in the future would add one column along the side of his or her matrix of character states giving a geological age for each taxon in the matrix (age is almost always known to good approximation from independent evidence), it would open the possibility of a host of new stratocladistic and stratophenetic approaches to phylogeny that tell us more than character analysis alone ever will, even for large numbers of taxa. Time, after all, is the independent variable in evolution, and time is the dimension biologists expect palaeontologists to bring to our common endeavour. □

Philip Gingerich is in the Museum of Paleontology, University of Michigan, Ann Arbor, Michigan 48109-1079, USA.

Of stars and stones

Alvar Ellegård

Records in Stone: Papers in Memory of Alexander Thom. Edited by C. L. N. Ruggles. Cambridge University Press: 1988. Pp.519. £50, \$100.

ARCHAEOASTRONOMY, the study of prehistoric peoples' star lore, stands to the history of astronomy as archaeology does to history. Everybody knows that Stonehenge has astronomical significance in so far as its main axis points in the direction of the rising midsummer Sun. That is a typical archaeoastronomical datum. The only practicable way for a non-technological civilization to make repeatable observations of the exact locations and movements of the heavenly bodies is to use the horizon as reference, and there is evidence from all over the world that illiterate peoples do indeed use it. For instance, some American Indians establish their annual calendar by marking the course of the Sun's rising point from a southerly extreme at the winter solstice to a northerly one in summer. Landmarks along the course may indicate important seasonal events.

The orientation of Stonehenge to the summer sunrise, and that of Newgrange, in Ireland, towards the winter sunrise, are exact enough for what was their (presumably) ritual purpose. But neither can indicate the exact day of the solstice, because (as the name itself implies) the sunrise point is nearly stationary around the solstices. Hence one of the crucial questions for the archaeoastronomer is, "Have we any evidence that the Neolithic Britons made more precise observations?". This question was hardly raised among British archaeologists until a professor of engineering science at Oxford, Alexander Thom, claimed in a book published in 1967 that the lay-out of stone rings and stone rows, and the orientation of single standing stones, proved that the ancient Britons did just this, several millennia BC.

The general public was fascinated by Thom's picture of the ancient Britons possessing a knowledge of astronomy at least as advanced as that of the ancient Egyptians or Chaldeans. A further boost to this way of thinking was provided by the considerable extension of the prehistoric time scale that took place in archaeology at about the same time, as a result of a calibration of carbon-14 dates by means of tree-ring datings. One of the consequences of this chronological revolution

was that diffusionist models of cultural change lost ground to theories of independent, indigenous development.

Thom arrived on the scene at the right moment. Television programmes were shown world-wide, popularizing his theories. We saw the Sun appearing for a brief second, setting or rising in a 'notch' in a distant mountain range, said to be unmistakably indicated by the main axis of a stone ring on a bleak hillside, or by an alignment of standing stones, or simply by the flat side of an impressive menhir. As sightlines of a length of 10–20 miles, directed at a particular point of the Sun's (or Moon's) disk, yield an accuracy of the order of minutes of arc, it was possible to claim that the megalith builders could in fact use their observations not only to

1985), in recognition of the decisive part he played in opening up this area of research. The 21 articles provide a balanced overview of current opinions in the field. It is interesting to see that none of the contributors try to support Thom's most extreme claims, to the effect that the stones should yield evidence of observations of the kind necessary, for instance, for the prediction of eclipses. The emphasis is now on placing the undoubted astronomical lore of the megalith builders in their social and cultural context. Two American contributors to the volume, A. Aveni and Ed Krupp, illustrate this trend very clearly. That is natural, because the American sites are in general much more recent, and at least in Mexico the buildings are far more elaborate than those that have been found in Neolithic Britain.

Chris Jennings

The centre-piece of the volume is the editor's own paper. Ruggles is the right man for this job. Having re-surveyed and re-analysed a central part of Thom's material, Ruggles's conclusion is definite: there is indeed a clustering of rough orientations towards the midwinter sunrise, and also, more remarkably (and hardly suspected before Thom), towards the extreme rising points of the Moon — what Thom called the major and minor standstills. However there is no valid evidence of any orientations of greater accuracy than one degree or so, for which only rather casual observations are needed. Apparently Thom had been misled by a quite natural, unintentional bias in his collection of data.

Ruggles rightly emphasizes the necessity of a strict methodology when using a statistical approach in archaeoastronomy. On the other hand, full account must also be taken of our background knowledge (or theories, or beliefs) of the culture concerned. There is little doubt that stone rings and



Mute witnesses — standing stones at Ballochroy, Kintyre, Scotland, one of the classic sites for students of prehistoric astronomy.

construct accurate calendars, but also to predict eclipses of the Sun and the Moon.

Partly because of Thom's popular success, the archaeologists, who at first had been rather hesitant, had to take his ideas seriously. This became even more evident when one archaeologist, Euan MacKie (also a contributor to the present volume), was able to produce from a Scottish site some hard archaeological evidence, predicted on the basis of Thom's theories. MacKie drew up a picture of Stone Age Britain in which there was an élite leisured class of specialists who had developed a sophisticated knowledge of astronomy.

Records in Stone is dedicated to the memory of Alexander Thom (who died in

menhirs were set out for many different purposes, and that such purposes changed over time. Lumping them all together will necessarily obscure much information contained in the data. The problem is to use our background knowledge without introducing a bias that works havoc with our (indispensable) statistical argument. This is a difficult task, but one that will surely be well worth the effort. There is substance in archaeoastronomy, although not quite as much as Thom wished to believe. □

Alvar Ellegård, Rattgatan 2, 42176 Frölunda, Sweden, was formerly a Professor in the Department of English at the University of Göteborg.

Limiting factors for modellers

J. P. Grime

Plant Strategies and the Dynamics and Structure of Plant Communities. By David Tilman. Princeton University Press:1988. Pp.360. Hbk \$45; pbk \$15.95.

THOSE attempting to model the full richness of vegetation dynamics must grapple with many facets of ecological research. They must recognize that each relevant field has its own concepts, and linkages must be established between some traditionally disparate fragments of the 'ecological jigsaw'. All embarking on such a venture must brace themselves against the natural reactions of specialists who see their work simplified in the cause of generality.

A further hazard is that the author of such an ambitious model is likely to introduce an unconscious bias promoting ideas derived from his own research experience and education. David Tilman has not escaped from this problem, and the strengths and weaknesses of his latest book are closely related to his origins as a modeller of the Princeton school and as a student of phytoplankton. In *ALLOCATE*, the simulation model which lies at the heart of his analysis, we find a bold attempt to identify the selection forces and constraints which dictate ecological specialization in plants. The result is an admirable but seriously flawed attempt to replace word models with mathematics and to test predictions against experimental data.

Tilman believes that competition is almost universally important as a determinant of ecological success, and he explains spatial variation in vegetation in terms of the differing abilities of species to tolerate particular resource shortages. The notion of relative tolerance as the mechanism of competition is contentious and at odds with the views of both Charles Darwin (as expressed in *The Origin of the Species*) and A. Milne (in *Mechanisms in Biological Competition*, edited by F.L. Milthorpe; Cambridge University Press, 1961). The origins of Tilman's model are readily apparent in his studies of the behaviour of mixed phytoplankton populations, where the struggle for existence frequently involves relative tolerance of resource shortages. Setting aside the vexed question of whether relative tolerance can be equated with competition, there are difficulties in extrapolating the model to terrestrial vegetation, in that equilibration time must be extended from days (phytoplankton) to more than 100 years (forest succession).

Tilman recognizes this limitation and introduces the notion of "transient dynamics" to describe the floristic changes

which may occur over many years before equilibrium is reached; this is a major concession and encourages the hope that *ALLOCATE* will evolve further towards ecological reality. In its present form, however, the model is severely restricted in general relevance, not only by its dependence upon equilibrium conditions of resource limitation, but by the prime importance which is given to the partitioning of resources between root and shoot. This reflects the author's assumptions about the availability of light and soil resources in various habitats. Tilman notes the conspicuous shaded stratum of dense productive vegetation and interprets this as evidence of light-limitation. But he fails to acknowledge the large body of literature which documents less obvious but equally important factors, such as the expanding zones of water and nutrient exhaustion which accompany root activity during each growing season. This patchiness and the attendant foraging responses of the roots of the fast-growing plants of

fertile soils, studied by M.C. Drew and others, strongly suggest that competitive ability at its most potent is an interactive property in which high rates of capture of all resources are sustained by simultaneous foraging responses which enable escape from depletion zones both above and below ground.

Tilman begins with the stated aim of exploring "the mechanisms of competition for soil resources and light". It is therefore disappointing to find that his model's dependence upon equilibration has necessitated a definition of competition which merges the process with a multiplicity of other factors determining the relative fitness of neighbouring populations. This is clearly unsatisfactory. As Milne observed, "competition ought to have only one meaning — clear, precise and unambiguous". □

J.P. Grime is a Professor in the Natural Environment Research Council Unit of Comparative Plant Ecology, University of Sheffield, Sheffield S10 2TN, UK.

The real thing

Murray Stewart

Artifacts in Biological Electron Microscopy. Edited by Richard F. E. Crang and Karen L. Klomparens. Plenum:1988. Pp.233. In the United States \$45, elsewhere \$54.

READING a book on artefacts in electron microscopy is rather like reading the *Karma Sutra*: one begins in a frenzy of anticipation and ends with a sense of amazement as to what people can do. In both instances one is presented with an intriguing catalogue of techniques that is useful for both instruction and inspiration, but that also contains quite a bit of information that may not be all that relevant in everyday life. And there is just so much that one is interested in that is not discussed at all.

Many chapters are excellent. Considerable attention is devoted to the difficulties encountered in sectioned embedded material; given the large number of cell biologists who now use these methods without being aware of the pitfalls that were common knowledge in the 1960s and 1970s, the inclusion of this topic was particularly appropriate. The chapters on scanning microscopy pinpoint many of the problems that crop up and describe ways of overcoming them. There is, too, a chapter on photography. Although the author's purpose is to provide advice as to how to record and print micrographs, this is nonetheless a fascinating article.

Other contributions are less worthwhile. The treatment of instrument-induced artefacts in transmission microscopy is rudimentary, and there is only the

most cursory account (often misleading if not actually incorrect) of image-formation mechanisms and focus-related artefacts. There is no real consideration of the compromise between resolution and contrast that has to be made with most biological material, or of the key importance of radiation damage in recording finer features or in making 3-D reconstructions from tilt series. The treatment of electron probe X-ray microanalysis is also too superficial to be really useful.

Quite astonishingly there is nothing on artefacts associated with negative staining or shadowing, both of which are extremely widely used, while most modern methods are either ignored or mentioned in asides. Such an important new technique as cryo-electron microscopy is discussed in terms of frozen sections, the extensive use of this method for observing material frozen in films of vitreous ice (an extremely fertile source of artefacts) being dealt with only in passing. Moreover, image processing, which is so powerful in obtaining reliable information especially from regular objects, is not included at all. That is a great pity — a discussion of artefacts in image processing of electron micrographs really would be something.

So, perhaps rather like the *Karma Sutra*, this book leaves one's interest aroused rather than satisfied, and wishing that, in places, it had been more explicit and that greater emphasis had been put on more up-to-date techniques. It nonetheless provides instructive reading for cell biologists, particularly those who only occasionally use electron microscopy. □

Murray Stewart is in the Division of Structural Studies, Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Information value of the rodent bioassay

Lester B. Lave, Fanny K. Ennever, Herbert S. Rosenkranz and Gilbert S. Omenn

Tests for human carcinogens using lifetime rodent bioassays are expensive, time-consuming and give uncertain results. For most chemicals such tests are not cost-effective.

In the past chemicals were implicitly tested in human beings, their toxicity being discovered through observing death and illness. Since the International Agency for Research on Cancer (IARC) lists only 26 chemicals or groups of chemicals as showing definite evidence of human carcinogenicity¹, either few chemicals of the 60,000 commonly used are carcinogens or few of the carcinogens are potent enough and reach enough people to have been recognized in humans. A preventive approach requires that chemicals must be tested before they are manufactured and marketed. Rodent lifetime bioassays have identified as carcinogens about 65% of 800 chemicals tested^{2,3}, an implausibly high proportion. Although scientists and regulators have reservations about using the lifetime rodent bioassay as the 'gold standard' for predicting human carcinogenicity, no generally accepted alternative exists^{4,5}.

In previous articles we developed a framework to measure the information content of the results of short-term, *in vivo* and *in vitro* tests used to classify chemicals⁶⁻⁹. This framework encompasses the cost of testing, the rates and social costs of misclassification, and the alternative of classification without testing. Here we apply this framework to the lifetime rodent bioassay. We find that the bioassay often does not provide information commensurate with its cost, implying that the regulatory policies of industrialized countries need to be changed.

The rodent bioassay has yielded positive results for all known human carcinogens, even benzene and arsenic (well-established human carcinogens, for which animal models were not available until recently). On this limited basis, the rodent bioassay appears to be 100% sensitive. But the usefulness of a test also depends on specificity (the proportion of true non-carcinogens found to be negative). The concordance of results between mice and rats can be used to infer the concordance between rodents and humans.

The US National Toxicology Program's studies find that 50% of 214 chemicals are positive in each species, but rats and mice gave concordant results (both positive or both negative) for only 70% of these chemicals¹⁰. Similar results have been obtained by others^{3,11}. Table 1 shows a representative matrix for 100 chemicals. Attributes of carcinogenicity such as

potency, route of administration, site of tumour, histopathology, strength of evidence, pharmacokinetics and extent of malignancy are ignored in our analysis. This positive-negative simplification, and the assumption that any chemical carcinogenic in mammals is carcinogenic in humans, do not reflect the growing sophistication of current mechanistic research in health assessments¹². But neither the regulatory agencies in the United States nor IARC use such data in their decisions, so our framework reflects current regulatory policy. More importantly, the framework provides the information basis for classifying potentially carcinogenic chemicals so as to minimize the social cost of misclassification.

How many chemicals that cause tumours in rodents are human carcinogens? Extrapolating from one species to another is fraught with uncertainty. Nevertheless, rats and mice are more similar biologically to each other than either is to humans¹³. Thus the 70% concordance in rodents appears to be an upper bound for rodent-human concordance; it might be much smaller. For example, Ennever *et al.*¹⁴ found that 19 out of 20 probable human non-carcinogens were positive in rodent bioassays, implying that specificity might be as low as 0.05.

According to current policy, a chemical causing tumours in a rodent species is suspect in humans, making 65 of the 100 chemicals in Table 1 suspect as human carcinogens. Table 2 results from assuming that 65% of chemicals are human carcinogens, and that concordance between rodents and humans is 0.7. In Table 2 rodent results correctly predict 50 of the 65 human carcinogens (77% sensitivity) and 20 of the 35 human non-carcinogens (57% specificity).

Few scientists believe that 65% of

chemicals in widespread use are carcinogens^{15,16}. Only 12% of chemicals were positive in systematic *Salmonella* mutagenicity testing of 2,205 chemicals in Japan¹⁷, and we postulate that no more than 10% of chemicals are carcinogenic to humans (from various mechanisms). If the concordance, sensitivity and specificity are 70% for rodent-human extrapolation, and 10% of chemicals are carcinogenic, Table 3 results: 30% of chemicals are classified inaccurately, with 3% false negatives (FN) and 27% false positives (FP).

Before the introduction of toxicity tests for chemicals, all 10 carcinogens (per 100 chemicals) (Table 3) were undetected initially and were treated as if they were FNs; there could be no FPs. The opposite extreme would classify all chemicals as carcinogens, again without testing, resulting in 90 FPs, with no FNs. Testing chemicals has a social cost equal to the sum of the direct costs of a rodent bioassay (\$1 million per chemical), plus the cost of 3% FNs and 27% FPs. We represent the social cost of an FP as x and of an FN as bx (refs 6-9) (see discussion below). For 100 chemicals, the cost of classifying all chemicals in Table 3 as non-carcinogens is $10bx$, of classifying all chemicals as carcinogens is $90x$, and of testing all the chemicals is $100 + 3bx + 27x$ (the units are millions of dollars).

Clearly, it is in everyone's interest to minimize the sum of the costs of testing and of mistakes in classification. Short-term tests and structure-activity analysis give a preliminary indication of the likelihood a chemical is a carcinogen. Then, for some values of b and x , a rodent bioassay would be best; for other values, no further testing is warranted. For example, suppose short-term tests predict a 65% likelihood that a chemical is a carcinogen (Table 2)

Table 1 Observed concordance between mouse and rat in the National Toxicology Program's lifetime bioassays

	Rat		
	+	-	
Mouse	+ 35	15	Concordance: agreement of test outcomes in both rodent species $(35 + 35)/100 = 0.70$
	- 15	35	Sensitivity: proportion of rat carcinogens that are positive in mice $35/(35 + 15) = 0.70$
			Specificity: proportion of rat non-carcinogens that are negative in mice $35/(15 + 35) = 0.70$
			%rodent carcinogens: ratio of all positive test outcomes to all chemicals $35 + 15 + 15 = 65$

+ and - indicate test outcome; that is, whether the chemical is positive or negative in the lifetime rodent bioassay.

and its $x = 1$ and $b = 10$. For 100 such chemicals, classifying all of them as carcinogens without further testing has a social cost of \$35 million, subjecting each to a rodent bioassay (with its 15 FNs and 15 FPs) would cost \$265 million, while classifying all chemicals as non-carcinogens without further testing would cost \$650 million. For the parameter values shown in Table 3, the social costs would be \$90 million, \$157 million and \$100 million, respectively.

The cost of an FP (x) is the social cost of not being able to use a chemical because it was erroneously thought to be a carcinogen (or the cost of labelling or of controls to limit human exposure). If profit lost by banning a chemical is the dominant factor, $x = 3$ implies a chemical would be generating \$600,000 in profits each year with future profits discounted at a rate of 20% per year. The social cost of an FN (bx) is the number of additional cancers due to unknowingly exposing people to a carcinogen. If $x = 3$ and $b = 10$, the social cost of proscribing a harmless chemical is \$3 million, while that of mistakenly exonerating a carcinogen is \$30 million.

When 10% of chemicals are carcinogens, with test accuracy shown in Table 3, $x = 3$ and $b = 10$, the cost of the three FNs is $3bx = \$90$ million; the 27 FPs would cost $27x = \$81$ million. The social cost of testing the 100 chemicals with a rodent bioassay would be $100 + 90 + 81 = \$271$ million.

Figure 1a-d shows the conditions under which testing is preferred, given the range of possible values for concordance, sensitivity and proportion of human carcinogens. In Fig. 1a, b and d , short-term tests are assumed to predict that a chemical has a 10% likelihood of being a carcinogen; in Fig. 1c, the likelihood is taken to be 2%, 10% or 50%. In Fig. 1a sensitivity, specificity and concordance are 70%; Fig. 1b varies concordance between 50% and 90%; Fig. 1d varies sensitivity from 50% to 90%. For values of x and b above the curve (diagonally shaded area in Fig. 1a), rodent bioassay testing minimizes social cost. For values of x and b below the curve and to the left (horizontally shaded area), chemicals should be classified as non-carcinogens without lifetime testing. For values of x and b below the curve and to the right (vertically shaded area), chemicals should be classified as carcinogens without lifetime testing.

In Fig. 1b, as concordance increases, there is a greater range of values of x and b that justify rodent bioassay testing. In Fig. 1c, rodent bioassay testing almost never has a lower cost than classifying chemicals as non-carcinogens without lifetime testing, when only 2% of chemicals are carcinogens. When 50% of chemicals are carcinogens, rodent bioassay testing almost never has a lower cost than classi-

Table 2 Implications of assuming concordance between rodent and human is 0.7, when the proportion of human carcinogens is 65%

	Humans		
	+	-	
	50	15	
	15	20	
Rodents			Concordance = 0.70 Sensitivity = 0.77 Specificity = 0.57 % human carcinogens = 65

Table 3 Misclassification when 10% of chemicals are assumed to be human carcinogens, and concordance, sensitivity and specificity are assumed to be 0.7

	Humans		
	+	-	
	7	27	
	3	63	
Rodents			Concordance = 0.70 Sensitivity = 0.70 Specificity = 0.70 % human carcinogens = 10

fying chemicals as carcinogens without lifetime testing. In Fig. 1d, the greater the sensitivity, the greater range of values of x and b implies that rodent bioassay testing has the lowest cost.

Only a few thousand of the 60,000 chemicals in use may be so valuable that the cost of an FP means that x is \$1 million or more¹⁸. The annual profits generated by a chemical, plus an estimate of the social loss by having to switch to a substitute, or the cost of stringent controls or labelling, could be estimated to determine the cost of an FP for any chemical (falsely classifying it as a carcinogen). The number of cancers and cancer deaths that might result from falsely classifying a carcinogen as safe (FN) could be estimated for any chemical by calculating the number of people exposed, the levels of exposure and the supposed potency of the carcinogen. Even if the estimates are uncertain, these two values and the likelihood that a chemical is a carcinogen (estimated from short-term tests) should be used to determine whether a chemical should be tested further in a rodent bioassay or classified on the basis of current information. This analysis indicates that few chemicals would be worth so uncertain and expensive a test as a rodent bioassay.

The chemicals selected for testing by

the National Toxicology Program are a distinctly non-random set, because 65% have been positive in at least one rodent species^{2,3,10,11}. If 65% of chemicals were human carcinogens (Table 2), most should be classified as carcinogens without a rodent bioassay: x must be greater than \$7 million and b between 0.3 and 1 to justify a rodent bioassay. Valuable process chemicals (intermediates and catalysts) might have such large values of x and small ratios of the cost of an FN to an FP but few others would.

What are the consequences when an innocuous chemical is falsely classified as a carcinogen (an FP)? A positive test result would mean that most chemicals would never be marketed in developed countries. Some of them would be of high value to society, particularly in replacing carcinogens, for example sodium cyclamate which could have replaced saccharin had it not been banned as a potential carcinogen^{19,20}. Valuable pharmaceuticals, pesticides or industrial chemicals might not be produced. A chemical which is known to be highly valuable and does not have a close substitute (for example methylene chloride), or is an intermediate kept entirely within the production process (for example styrene), might be exempted from a ban. Even so, the costs of engineering controls to minimize human exposure could be hundreds of millions of dollars.

Falsely classifying a chemical as a non-carcinogen (an FN) exposes people to a carcinogen. The social cost will depend on the number and types of cancers that would result, which are a product of the level of exposure, the number of people exposed, the potency of the chemical and the malignancy of the tumour.

The monetary value of preventing a premature death from cancer can be estimated from current US government criteria^{21,22}. While society seeks to prevent cancers, we stop short of doing everything possible. Devoting more resources to helping people to give up smoking or change their diet, or to screening for breast and colon cancer, would save many lives²³. Since all tests have FNs, preventing all cancers caused by chemicals would require banning all potential chemical carcinogens. Rather, society has chosen

1. IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs 1 to 42 Suppl. 7 (International Agency for Research on Cancer, Lyon, 1987).
2. Gold, L.S. et al. *Environ. Hlth Perspect.* **58**, 9-319 (1984).
3. Gold, L.S. et al. *Environ. Hlth Perspect.* **67**, 161-200 (1986).
4. Clayton, D.B. *Mut. Res.* **185**, 243-269 (1987).
5. Ames, B.N., Magaw, R. & Gold, L.S. *Science* **236**, 271-280 (1987).
6. Lave, L.B., Omenn, G.S., Heffernan, K.D. & Dranoff, G. *J. Am. Coll. Tox.* **2**, 125-130 (1983).
7. Lave, L.B. & Omenn, G.S. *Nature* **324**, 29-34 (1986).
8. Lave, L.B. & Omenn, G.S. *J. Am. Coll. Tox.* **7**(5), 565-574 (1988).
9. Omenn, G.S. & Lave, L.B. *Mut. Res.* **205**, 41-49 (1988).
10. Zeiger, E. *Cancer Res.* **47**, 1287-1296 (1987).
11. Tennant, R.W. et al. *Science* **236**, 933-941 (1987).
12. Ashby, J. & Tennant, R.W. *Mutagen. Res.* **204**, 17-115 (1987).
13. Hart, R.W. & Fishbein, L. in *Toxicological Risk Assessment Vol. 1: Biological and Statistical Criteria* (eds Clayton, D.B., Krewski D. & Munro I). 3-40 (CRC Press, Boca Raton, Florida, 1985).
14. Ennever, F.K., Noonan, T.J. & Rosenkranz, H.S. *Mutagenesis* **2**, 73-78 (1987).
15. Huff, J.E., McConnell, E.E. & Haseman, J.K. *Environ. Mutagen.* **7**, 427 (1985).
16. Rosenkranz, H.S. *Environ. Mutagen.* **7**, 428 (1985).
17. Matsushima, T. *Chemical Safety Evaluation in Japan* 175-178 (1987).
18. *Chemical and Engineering News* 24, 8 June (1987).
19. National Research Council *Saccharin: Technical Assessment of Risks and Benefits* (National Academy Press, Washington, DC, 1978).
20. National Research Council *Evaluation of Cyclamate for Carcinogenicity* (National Academy Press, Washington, DC, 1978).

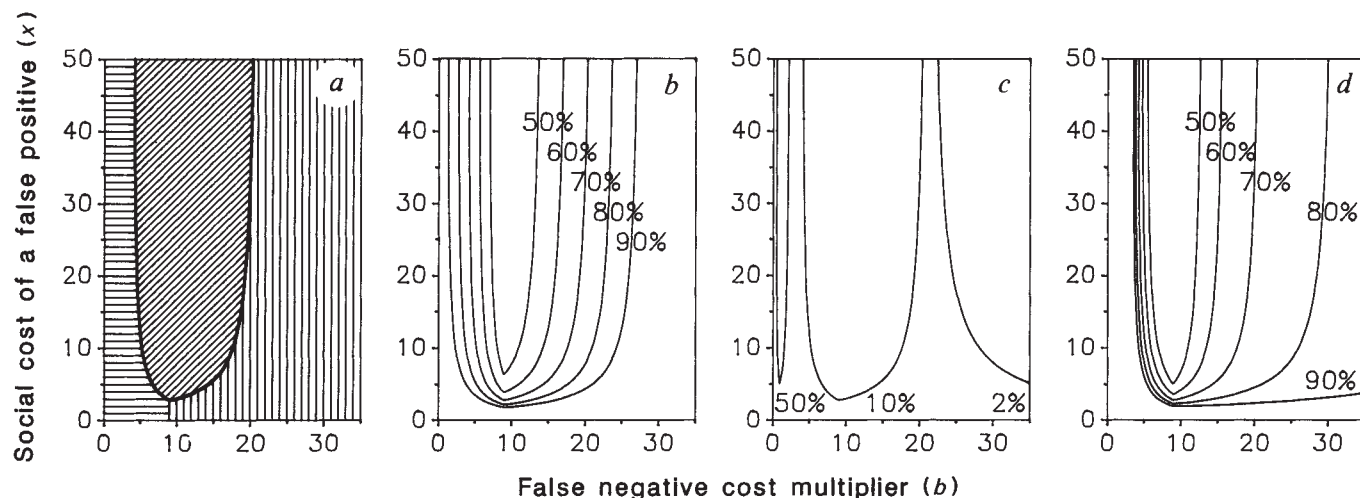


Fig. 1 Critical values of x and b that justify carcinogenicity testing with a lifetime rodent bioassay.

The vertical axis (x) is the social cost of a false positive (FP) — the profit lost by banning a chemical, cost of a substitute chemical or cost of imposing stringent controls or labelling to reduce human exposure. The horizontal axis (b) is the ratio of the social costs of a false negative (FN) to those of an FP — the ratio of the cost of additional cancers from exonerating a carcinogen to the loss due to indicting an innocuous chemical. Values of x and b lying above the curve (diagonally shaded area in Fig. 1a) indicate that a rodent bioassay, and associated misclassification, costs less than classifying a chemical without further testing (with the associated misclassification costs). Values of x and b below the curve and to the left (horizontal shading in Fig. 1a) indicate that classifying the chemical as a non-carcinogen without further lifetime testing has the lowest social cost. Values of x and b below the curve and to the right (vertical shading in Fig. 1a) indicate classifying the chemical as a carcinogen without further testing has the lowest social cost.

This curve and interpretation of the three areas are repeated in Fig. 1b–d. Figure 1a assumes that 10% of chemicals are human carcinogens (or that a short-term test or structure–activity analysis indicate that there is a 0.1 likelihood this chemical is a carcinogen). The rodent bioassay is assumed to cost \$1 million, to have 70% concordance (rodent test results agree with human results for 70% of chemicals), 70% sensitivity (70% of human carcinogens are positive in rodent bioassays) and 70% specificity (70% of human non-carcinogens are negative in rodent bioassays). Figure 1b repeats Fig. 1a, but varies concordance from 50% to 90%. Figure 1c repeats Fig. 1a, but allows the proportion of chemicals that are carcinogens to be 2%, 10% or 50%; the likelihood that a chemical is a carcinogen (2%, 10% or 50%) is derived from short-term tests or a structure–activity analysis. Figure 1d repeats Fig. 1a, but varies sensitivity from 50% to 90%.

not to do everything possible, or even everything that is moderately expensive, to prevent cancer.

However, relatively few chemicals are used so as to expose millions of people to them; few exposures are so high and few chemicals so potent that thousands or even dozens of cancers would be expected each year. The benefits of chemicals classified as carcinogens could be great enough to result in their continued use. For example, some of the drugs used in chemotherapy are carcinogens. Also, amidst uncertainty about whether saccharin really is a human carcinogen, many people continued to use it in the belief that the benefits outweighed the risks.

Thus assuming that the concordance between rats and mice for carcinogenicity provides an upper bound for the concordance between rodents and humans has important implications for testing strategies. For all but a few chemicals, perform-

ing a lifetime rodent bioassay has a higher social cost than classifying on the basis of short-term test results and estimates of the cost of an FP and FN. If only 10% of chemicals are human carcinogens, the lifetime bioassay should be used only for chemicals of high value that might potentially cause many cancers. If concordance or sensitivity were below 70%, or if the likelihood of a chemical being a carcinogen is below 10%, rodent bioassays become less attractive than classifying chemicals as non-carcinogens without further testing. Similarly, if the likelihood that a chemical is a human carcinogen is 50% or more, a rodent bioassay is more costly than classifying the chemical as a carcinogen without further testing.

For almost all of the chemicals tested to date, rodent bioassays have not been cost-effective. They give limited and uncertain information on carcinogenicity, generally give no indication of mechanism of action, and require years to complete. Instead, some of these resources should be devoted to improving the sensitivity, specificity and cost-effectiveness of alternatives to the long-term bioassay, such as *in vitro* and *in vivo* short-term test schemes^{6–9, 24–27}. Fortunately, the rigorously conducted bioassay testing to date, particularly by the National Toxicology Program, has been used also to evaluate short-term tests and hypotheses about structure–activity relationships²⁸ and provide potency estimates for risk analysis.

The value-of-information analysis pre-

sented here sets out the detailed conditions under which a rodent bioassay should be used for carcinogenicity testing. Regulators should estimate the values of x (from company profits and the social cost of substituting the next best chemical, labelling or the cost of control); bx (from estimating the number of cancers that might result from exonerating this chemical); and proportion of human carcinogens (via structure–activity analysis and short-term tests) for each chemical. These formulae then indicate whether the chemical should be tested in a rodent bioassay or classified without further testing as a non-carcinogen or as a carcinogen.

Our framework reveals that the lifetime rodent bioassay is not cost-effective for most chemicals. This assay is a valuable tool for investigating pharmacokinetics, mechanisms of action and differences among species. But it is rarely the best approach for deciding whether to classify a chemical as a human carcinogen. □

We thank Mortimer Mendelsohn and his colleagues, Robert Moolenaar, Fred Hoerger and Ernest Weiler for comments on this article.

Lester B. Lave is in the Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, USA. Fanny K. Ennever and Herbert S. Rosenkranz are in the Department of Environmental Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA. Gilbert S. Omenn is in the School of Public Health and Community Medicine, and Dana Program in Genetics and Environmental Health, University of Washington, Seattle, Washington 98195, USA.

DC, 1985).

21. Lave, L.B. *The Strategy of Social Regulation* 39–41 (Brookings Institution, Washington, DC, 1981).
22. Mendeloff, J.M. *The Dilemma of Toxic Substance Regulation* 21–52 (MIT Press, Cambridge, Massachusetts, 1988).
23. US National Cancer Institute *Cancer Goals for 2000*, (Department of Health and Human Services, Washington, DC, 1987).
24. Chankong, V., Haimes, Y.Y., Rosenkranz, H.S. & Pet-Edwards, J. *Mut. Res.* **153**, 135–166 (1985).
25. Ennever, F.K. & Rosenkranz, H.S. *Mutagenesis* **1**, 293–298 (1986).
26. Rosenkranz, H.S., Klopman, G., Chankong, V., Pet-Edwards, J. & Haimes, Y.Y. *Environ. Mutagen.* **6**, 231–258 (1984).
27. Weisburger, J.H. & Williams, G.M. *Science* **214**, 401–404 (1981).
28. Frierson, M., Klopman, G. & Rosenkranz, H.S. *Environ. Mutagen.* **8**, 283–327 (1986).

Solar oscillation frequencies and the equation of state

J. Christensen-Dalsgaard^{*§}, W. Däppen^{†‡} & Y. Lebreton[†]

^{*} Astronomisk Institute, Aarhus Universitet, DK-8000 Aarhus C, Denmark

[†] Observatoire de Paris-Meudon, F-92195 Meudon, France

[‡] Service d'Astrophysique, CEN Saclay, F-91191 Gif-sur-Yvette, France

Observed oscillation frequencies of the Sun can be used to investigate the properties of matter under conditions that cannot be achieved on Earth. In particular the frequencies are sensitive to the equation of state. A recently developed treatment of the partition functions leads to a substantial improvement in the agreement between the observed and the computed frequencies.

ONE of the goals of helioseismology is to use the observed frequencies to probe the properties of matter under solar conditions. Here we study the sensitivity of the frequencies to the equation of state used in the computation of the solar model. Mihalas, Hummer and Däppen¹⁻³ (hereafter MHD) have recently developed a new treatment of the equation of state; this causes a significant change, relative to a model using a simpler equation of state, in the sound speed in the solar convection zone, which leads to a substantial improvement in the agreement between the observed and computed frequencies. The frequency differences between the models are simply related to the change in the sound speed, which in turn can be understood in terms of modifications to the partition functions. Thus, there appears to be a direct relation between a microscopic property of the gas, namely the partition function, and the observable frequencies. This demonstrates the power of helioseismology to address fundamental physical questions. Furthermore, unless the computed model suffers from other errors, the MHD equation of state appears to provide a good representation of thermodynamic conditions in the solar convection zone.

The MHD equation of state has been developed as part of a continuing opacity recomputation project⁴. Although the basic concept of the MHD equation of state is conventional, involving minimization of the free energy, it is characterized by detailed internal partition functions of a large number of atomic, molecular and ionic species. Less conventional equations of state (which are often associated with the Planck-Larkin partition function) are currently being developed^{5,6} and their predictions could well be different from those of MHD. Laboratory spectra have been reproduced successfully using the MHD equation of state⁷. However, in view of the important open questions regarding the equation of state, more detailed comparisons of the predictions of representative equations of state with observational data are needed. The observed frequencies of solar oscillations constitute an excellent data set for this purpose. Here we describe a first attempt at a comparison which relates the equation of state to the properties of modes spanning the entire range of observed frequencies and horizontal scales in the solar five-minute band.

Models of solar structure, and the oscillation frequencies which they predict, depend on the assumed properties of matter in the Sun, as described by the equation of state, the opacity and the energy generation rate, as well as on the description of the turbulent convection near the solar surface and of the solar atmosphere. In addition, the initial chemical composition of the Sun must be specified. In most of the Sun, the oscillations can

be assumed to be adiabatic. Then their behaviour is controlled by the pressure p , density ρ and the adiabatic exponent $\Gamma_1 \equiv (\partial \ln p / \partial \ln \rho)_s$ (the derivative being at constant specific entropy s), as functions of distance r to the centre. The quantities p and ρ are related through the equation of hydrostatic support, and, given the composition, Γ_1 is determined from p and ρ by the equation of state. In general, however, the determination of $\rho(r)$ requires a consideration of the energy generation and transport, as well as the equation of state; thus, without additional assumptions it is difficult to use observed oscillation frequencies to study any single aspect of the physics.

The situation is considerably simplified, however, in most of the solar convection zone. Because of the relatively high thermal capacity of the gas, energy transport by convection requires that the actual temperature gradient in the Sun exceeds the adiabatic value by only a very small amount; that is, only a very small superadiabatic gradient is required. Thus, the stratification of density is also nearly adiabatic:

$$\frac{1}{\Gamma} \equiv \frac{d \ln \rho}{d \ln p} \approx \frac{1}{\Gamma_1} \quad (1)$$

To the extent that this equation is satisfied, the structure of the convection zone is specified by the (constant) value of s , the equation of state and the composition, which is constant throughout the convection zone because of the vigorous mixing. Only very near the surface is the thermal capacity so low that energy transport requires an appreciable superadiabatic gradient. The integral over this superadiabatic gradient fixes s . In most computations s is determined by selecting the value of a parameter in the description of the convective energy transport. Commonly used is the mixing-length prescription, where energy transport is assumed to occur through the motion of eddies which preserve their identity while moving a specified distance, after which they are dissolved in the surrounding medium; here the parameter is the ratio between that distance, called the mixing length, and the pressure scale height. The parameter characterizing convection, and the initial hydrogen abundance, are adjusted until the model has the correct radius and luminosity.

Sensitivity to solar structure

We consider only the five-minute oscillations, which correspond to acoustic modes of either high radial order or high degree. These modes 'feel' the structure of the Sun mainly through the adiabatic sound speed c , given by

$$c^2 = \frac{\Gamma_1 p}{\rho} \quad (2)$$

The modes extend between a lower turning point at $r = r_*$,

[§] Also at HAO/NCAR, PO Box 3000, Boulder, Colorado 80307, USA.

approximately determined by

$$\frac{2\pi\nu}{L} = \frac{c(r_i)}{r_i} \quad (3)$$

and an upper turning point just beneath the photosphere; here ν is the (cyclic) frequency of oscillation and $L = \sqrt{l(l+1)}$, where l is the degree of the mode. The condition (3) is equivalent to the vanishing of the vertical component of the wave vector in an asymptotic description of the oscillations as locally plane waves⁸, corresponding to a total internal reflection of the mode. Thus the modes are restricted to a layer beneath the surface whose depth decreases with increasing degree. Modes with $l \geq 50$ are essentially confined within the convection zone. Accordingly, the frequencies of such modes are determined predominantly by s , the composition and the equation of state.

The decrease in the density scale height near the surface causes the reflection at the upper turning point; it is located approximately at the point where the frequency is equal to the local acoustic cutoff frequency⁹ ω_c . This position is essentially independent of l , whereas its depth increases with decreasing frequency. Although ω_c depends on the uncertain detailed structure of the upper, significantly superadiabatic, part of the convection zone, it seems fairly certain that for $\nu < 2,000 \mu\text{Hz}$ the turning point is in a region where both the stratification and the oscillations are nearly adiabatic^{10,11}. Above the turning point the modes are evanescent, and the energy density decreases approximately exponentially. This property is reflected in the insensitivity of low-frequency modes to changes in the model very near the surface¹⁰; in addition, the small amplitude in the non-adiabatic region, relative to the amplitude in the bulk of the Sun, causes the rapid decrease in observed^{12,13} and computed^{14,15} oscillation line widths with decreasing frequency.

Changes in the structure of the model, or in the assumed physics of the oscillations, introduce a perturbation $\delta\nu_{nl}$ in the oscillation frequency of a mode of degree l and order n that depends on the overlap between the imposed changes and the eigenfunction of the mode. In addition, $\delta\nu_{nl}$ depends on the mode inertia:

$$E_{nl} = \int_0^R |\delta\mathbf{r}_{nl}|^2 \rho r^2 dr \quad (4)$$

where $\delta\mathbf{r}_{nl}$ is the displacement vector and the integral is from the centre to the surface radius R of the model. Loosely speaking, with increasing l the oscillations involve a smaller part of the Sun and have smaller inertia, and hence their frequencies are more sensitive to changes in the model. A more precise statement of this property results from a perturbation analysis of the oscillation equations¹⁶. To remove the part of the variation in $\delta\nu_{nl}$ that is caused by the dependence of inertia on l , it is convenient to consider the scaled frequency difference $Q_{nl}\delta\nu_{nl}$ with

$$Q_{nl} = \frac{E_{nl}}{\bar{E}_0(\nu_{nl})} \quad (5)$$

where $\bar{E}_0(\nu)$ is a fit to the inertia of the radial modes, and the ratio is at fixed photospheric amplitude. This effectively normalizes the frequency change to the value that would result in a radial mode of the same frequency, if the change in the model had been restricted to the region where the mode under consideration is trapped. As argued above, Q_{nl} decreases with increasing l at fixed frequency, although for $l \leq 20$ it is close to unity. For $l = 1,000$, $Q_{nl} \approx 0.1$.

Our ability to use the oscillation frequencies to probe the solar interior is limited by our inadequate understanding of the structure and dynamics of the upper part of the solar convection zone, and of the effects of this region on the physics of the oscillations, including their damping and excitation. Nevertheless, because such uncertainties are almost certainly confined to a very thin region where the modes propagate nearly vertically,

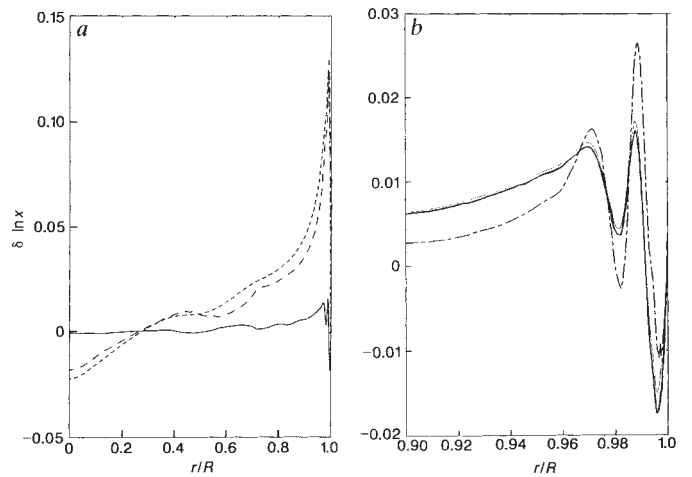


Fig. 1 Differences in variables x between the MHD and the EFF models, in the sense MHD – EFF, at fixed radius r . Illustrated are $\delta \ln p$ (short-dashed line), $\delta \ln \rho$ (long-dashed line), $\delta \ln c$ (heavy solid line) and $\delta \ln \Gamma_1$ (alternate short and long dashes). An expanded section of a is shown in b , for the region close to $r/R = 1.0$. The thin solid line in b shows $\delta \ln c$, calculated from equations (6) and (7), assuming that $\delta \ln \Gamma/\Gamma = \delta \ln \Gamma_1/\Gamma_1$.

their effects on the dynamics of the oscillations are largely independent of the horizontal wavenumber of the modes, although in general they do depend on the mode frequency. Thus, in terms of the scaled frequency differences, the uncertain effects of the superficial layers should be functions of frequency alone. Furthermore, modes with $\nu < 2,000 \mu\text{Hz}$ have little energy in these layers, and hence are probably essentially unaffected by them.

Model computations

To test the effects on the oscillation frequencies of modifications to the equation of state, we have computed a model of the present Sun using the MHD¹⁻³ equation of state. Whereas conventional equations of state, such as the EFF formulation discussed below, often assume that bound configurations (atoms, molecules, ions) are in their ground state, the MHD equation of state starts out from detailed internal partition functions. Occupation probability weights are assigned to the excited states such that the partition functions remain finite. Neutral and charged surrounding particles reduce the occupation probabilities of the upper levels, the first acting through an excluded volume effect, the second through Stark ionization. The energy levels themselves are assumed to be unperturbed and the continuum is not lowered. The Coulomb correction (treated in the Debye-Hückel approximation) is included in the MHD equation of state. Furthermore, its numerical realization has been made very smooth, accurate and reliable; this guarantees the consistency of the thermodynamical quantities which is needed in stellar pulsation calculations. Though the MHD equation of state was originally developed for plasmas under the conditions found in stellar envelopes, it can be applied to the entire Sun, because it takes into account the physical effects (partial electron degeneracy, Coulomb pressure, excluded-volume terms) that are usually present in standard calculations. For the model discussed below, a simplified mixture of chemical elements was used, consisting of H, He, O and Fe; the O and Fe abundances were adjusted to match the actual combined abundance of C, N, O, and of the remaining heavy elements, respectively. A separate calculation in which C, N and O were all included gave very similar results.

The model computed with the MHD equation of state was compared with a model that employed the much simpler, although still thermodynamically consistent, Eggleton, Faulkner and Flannery¹⁷ (EFF) formulation, which neglects the Coulomb

effect. Apart from the equation of state, the physics of the two calculations¹⁸ was the same. We used 601 spatial mesh points and 25 time steps, thus ensuring that numerical errors in the model calculation contributed less than $0.3 \mu\text{Hz}$ to the frequency error, scaled by Q_{nl} . The error introduced by the frequency calculation, given the model, was somewhat smaller. As usual the models were calibrated, by adjusting the composition and the parameter characterizing convection, to achieve the correct luminosity and radius.

Because the MHD equation of state is too complicated to incorporate directly into a stellar evolution programme, it was implemented through interpolation in previously computed tables on a mesh in $\log \rho$, $\log T$ and X , T being temperature and X the hydrogen abundance by mass. To test the effect of the interpolation, we computed a set of tables using the EFF formulation, on the same grid in $(\log \rho, \log T, X)$, and compared a model computed with interpolation in these tables with a model where the EFF equation of state was incorporated directly. The error introduced by the interpolation was less than 0.1% for the sound speed, and, except in the core of the model, was less than 0.2% for pressure and density. These errors are insignificant compared with the effects of changing the equation of state. To reduce the effects of the interpolation error further, the MHD results are compared with results obtained with interpolation in the EFF tables.

Figure 1 illustrates differences $\delta \ln p$, $\delta \ln \rho$, $\delta \ln c$ and $\delta \ln \Gamma_1$ between the MHD and the EFF models; here δ denotes differences at fixed r . The modification of the equation of state clearly has a substantial effect on pressure and density in the outer part of the convection zone. The change in the sound speed is smaller, but still significant. It should also be noted that to obtain the correct luminosity, the surface value of the hydrogen abundance, X , had to be 0.743 for the MHD model, whereas the value for the EFF model was 0.733. This change in X is probably caused by the tendency for the Coulomb effect to reduce pressure in the MHD model; increasing X partly compensates for this by decreasing the mean molecular weight¹⁹.

As shown in Fig. 1b, $\delta \ln c$ closely resembles $\delta \ln \Gamma_1$. From equation (2) it follows that

$$\delta \ln c = \frac{1}{2} \left(\delta \ln \Gamma_1 + \delta \ln \frac{p}{\rho} \right) \quad (6)$$

Also it may be shown from the equation of hydrostatic support, and the definition of Γ in equation (1), that

$$\delta \ln \frac{p}{\rho} = \left(\frac{r}{R} \right)^2 H_p^{-1} \left[H_{p,0} \left(\delta \ln \frac{p}{\rho} \right)_0 + \int_r^R \left(\frac{R'}{r'} \right)^2 \frac{\delta \ln \Gamma}{\Gamma} dr' \right] \quad (7)$$

Here H_p is the pressure scale height, and subscript zero indicates values at $r = R$. In deriving this equation the mass exterior to radius r was neglected, but the variation of gravity with radius was taken into account. In the present case the perturbation to the superadiabatic gradient is small; thus $\delta \ln \Gamma / \Gamma$ can be approximated by $\delta \ln \Gamma_1 / \Gamma_1$ in equation (7). Making this approximation, neglecting the surface term, and using equation (6) for $\delta \ln c$, we obtain the result shown as a thin solid line in Fig. 1b. It is evident that the change in the sound speed is essentially determined by $\delta \Gamma_1$.

The change in Γ_1 can be written as

$$\delta \ln \Gamma_1 = (\delta \ln \Gamma_1)^{(i)} + \left(\frac{\partial \ln \Gamma_1}{\partial \ln p} \right)_\rho \delta \ln p + \left(\frac{\partial \ln \Gamma_1}{\partial \ln \rho} \right)_p \delta \ln \rho \quad (8)$$

Here $(\delta \ln \Gamma_1)^{(i)}$ is the intrinsic change, at fixed p and ρ , brought about by the change in the equation of state or a change in

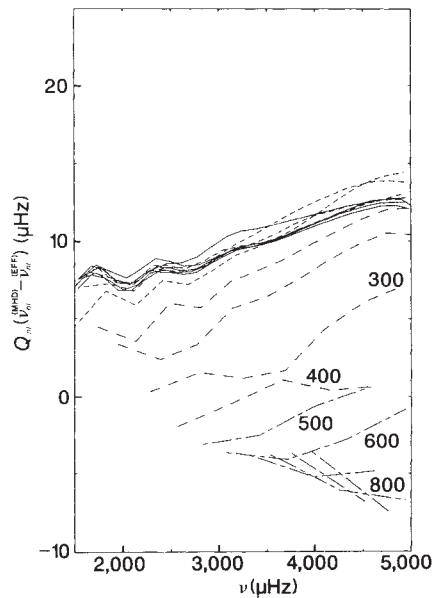


Fig. 2 Frequency differences $\nu_{nl}^{(\text{MHD})} - \nu_{nl}^{(\text{EFF})}$ between the MHD and the EFF models, for selected values of l . The differences have been scaled by the normalized mode inertia (compare equation (5)). Points corresponding to a given value of l have been connected as follows: $l = 0-30$ (solid line); $l = 40-100$ (short-dashed line); $l = 150-400$ (long-dashed line); and $l = 500-1,000$ (alternate short and long dashes). In addition a few values of l are shown in the figure.

composition; the remaining terms arise from the changes in p and ρ and the dependence of Γ_1 on these quantities. The contribution of these terms to the difference in Γ_1 between the MHD and the EFF models is small. Consequently, $\delta \ln \Gamma_1$, and hence $\delta \ln c$, can to a large extent be understood directly in terms of the change in the equation of state.

In this part of the Sun, the principal effect of the MHD equation of state is caused by the fact that the internal partition functions of H, He and He^+ are larger than usual, which increases the statistical weight of these species. These larger statistical weights lead to smaller ionization fractions than are obtained when all atoms and ions are in the ground state. In the partition function, the ground state (with (negative) energy E_0) dominates each excited state (with (negative) energies E_j) by a factor of $\exp[-(E_j - E_0)/kT]$, but the large number of excited states can lead to a sizeable correction to the total statistical weight of the bound system. The increase of the MHD partition function relative to that of the EFF case can be as high as 10-30%. This increase is then translated into a decrease of the H, He and He^+ ionization fractions. In a partially ionized plasma Γ_1 is lowered from its ideal value (at full ionization or recombination) of 5/3; the minimum values are ~ 1.20 and ~ 1.55 in the H-ionization zone and the He^+ -ionization zone, respectively. The effect of MHD is to push down the zone where this lowering of Γ_1 takes place. Smaller effects, namely the detailed temperature and density dependence of the internal partition functions, affect the minimum value of Γ_1 as a function of r ; the minimum value of Γ_1 in the H-ionization zone of MHD is seen to be slightly larger than that of EFF. In this way the large sinusoidal fluctuation of $\delta \Gamma_1$ in the region between 0.985 and 1.0 in Fig. 1b can be understood at least qualitatively. The effect of helium is more difficult to discuss, though in principle similar mechanisms apply.

Scaled frequency differences between the models are shown in Fig. 2. At low degree, $Q_{nl} \delta \nu_{nl}$ is positive and essentially independent of l . As l increases, the frequency difference decreases, becoming negative at the highest degrees considered. This behaviour can be readily understood in terms of the difference in sound speed. From a crude approximation to the asymptotic theory of acoustic oscillations²⁰, we find

$$\frac{\delta \nu}{\nu} \approx \frac{\int_{r_1}^R \frac{\delta c}{c} \frac{dr}{c}}{\int_{r_1}^R \frac{dr}{c}} \quad (9)$$

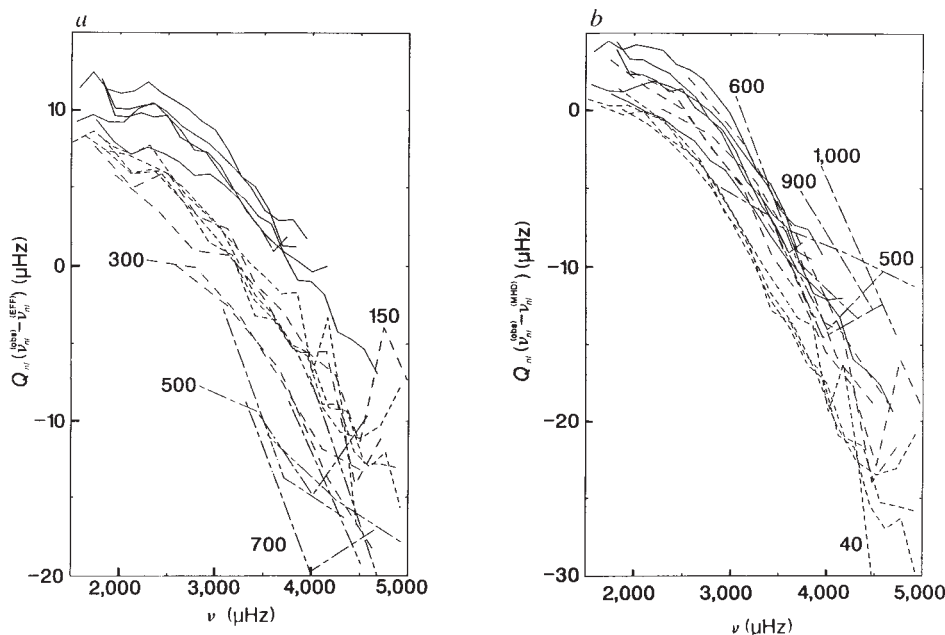


Fig. 3 Scaled differences $Q_{nl}(\nu_{nl}^{(\text{obs})} - \nu_{nl}^{(\text{EFF})})$ between observed frequencies and frequencies computed for the EFF model (a), and the corresponding scaled differences $Q_{nl}(\nu_{nl}^{(\text{obs})} - \nu_{nl}^{(\text{MHD})})$ for the MHD model (b). The key is the same as that for Fig. 2.

Here the denominator is roughly proportional to the mode inertia, so that the scaled frequency difference behaves approximately as the numerator in equation (9). If $\delta c/c$ is large mainly in a certain region of the model, it might therefore be expected that the scaled frequency difference is similar, and hence independent of l , for all modes that penetrate substantially beneath this region; modes with turning points outside the region should not be affected. As c decreases towards the surface, the behaviour of δc for $r/R > 0.9$ dominates $\delta \nu$. Therefore, for modes of degree less than ~ 100 , for which $r_l/R < 0.9$ in the frequency range considered, the scaled frequency difference is predominantly a function of ν . As l increases beyond 100, and r_l/R increases beyond 0.9, the contribution from the region where δc is positive is decreased, causing the decrease in $\delta \nu$ with increasing l . For $l \geq 500$, the region of negative δc very near the surface dominates, and $\delta \nu$ is negative.

Comparison with observed frequencies

It is of obvious interest to compare the computed frequencies with those observed. Figure 3a shows scaled differences between the observed frequencies in Duvall *et al.*²¹ and in Libbrecht and Kaufman²², and those computed for the EFF model. The quantity $\delta \nu$ shows a substantial dependence on the degree. This suggests²³ that a significant contribution to the frequency error comes from an extended part of the Sun, causing the frequency difference to depend on the position of the lower turning point. A similar dependence on degree is observed in the comparison in Fig. 2 between the MHD and the EFF models; here, however, the assumptions used in the convection zone and the atmosphere are essentially the same in the two cases, and hence the frequency-dependent part of $\delta \nu$ is absent. The presence of a substantial frequency error at low frequencies also indicates that the errors in the model extend beyond the region where non-adiabaticity and other explicitly neglected effects are significant. In contrast, the scaled frequency differences between the observation and the MHD model, shown in Fig. 3b, show little systematic variation with l , and the error at low frequency has been reduced substantially. Thus, it appears that much of the error in the interior of the model has been eliminated by using the MHD equation of state.

This conclusion is valid only to the extent that other sources of error can be neglected. As pointed out above, the principal quantities affecting the adiabatic part of the convection zone are the hydrogen abundance X and the specific entropy s . Both are in effect determined by requiring a continuous match

between the convection zone and the radiative interior, and hence are sensitive to errors in the physics of the latter. In fact, other model calculations, using newer and presumably more accurate opacity tables^{24–27}, have yielded values of X lower by about 0.04 than the one found here. The effect of such a change on the structure of the convection zone evidently depends on the associated change in s . Envelope calculations²⁸ have shown that, taken in isolation, a change in X of 0.04 results in a change in the sound speed of less than 0.5%, except in the outer 0.5% of the radius; this is substantially smaller than the effect on c obtained by changing the equation of state. Furthermore, given the overall agreement between the MHD model and those properties of the observed frequencies that are sensitive to the interior structure of the model, any modifications to X or s must be such as to result in a model whose sound speed is similar to that of the MHD model.

It should also be noticed that there remain systematic variations with l in the scaled frequency difference. The decrease in the difference as l is increased from 30 to 40 was discussed in ref. 23; it is related to the sound speed difference beneath the convection zone obtained by asymptotic inversion of the oscillation frequencies⁸. Furthermore, the differences for $l = 900$ and 1,000 appear to lie significantly above those for modes of lower degree. This could be due to remaining errors in the sound speed in the outer 0.5% of the model, possibly caused by an error in the hydrogen abundance, or it could be due to, for instance, l -dependent effects of convection. Systematic errors in the observations cannot, however, be ruled out.

It is desirable to identify the dominant sources of error in the computed frequencies, which, as we have argued, must be localized close to the solar surface. Non-adiabaticity generally reduces the computed frequencies^{29,30}; the frequency dependence of the effect is not unlike that seen in Fig. 3b, but the magnitude is probably insufficient to account for the total difference. Decreasing the magnitude of the superadiabatic gradient, but extending the region where it is significant (thereby keeping s fixed), tends to increase the oscillation frequencies¹⁰. In fact it is likely that the superadiabatic gradient should be larger and more narrowly confined than in the models computed here. In non-local mixing-length theory the eddies sense an average over the mixing length of the superadiabatic gradient; consequently the actual superadiabatic gradient is larger than in the local mixing-length theory used here³¹. Nonlinear granulation calculations support this result. The increase in the superadiabatic gradient might be expected to decrease the frequencies,

thus improving the agreement with observations. The inclusion of dynamical effects of convection, through the perturbation to the turbulent Reynolds stress, apparently tends to increase the computed frequencies³⁰. None of these effects have been studied in sufficient detail to enable definite conclusions to be drawn about their likely contributions to the observed frequency differences. It is interesting, however, that a recent fully compressible nonlinear calculation of solar granulation³² showed oscillations whose frequencies deviated from linear adiabatic frequencies of a similar normal model by an amount that, when properly scaled, was not dissimilar to the difference shown in Fig. 3b. Thus, it appears that with an adequate description of the equation of state and a detailed model of the upper part of the convection zone, we may be able to account for the gross

features of the discrepancy between observed and computed frequencies for the solar five-minute oscillations.

We are grateful to D. O. Gough, D. G. Hummer, D. Mihalas, Å. Nordlund and M. J. Thompson for useful discussions, and to R. L. Gilliland, D. O. Gough and M. J. Thompson for reading and commenting on an early version of the paper. J. Toomre is thanked for providing access to the computing resources of the University of Colorado during the Helioseismology Workshop at the Joint Institute for Laboratory Astrophysics, July 1988. Computing was also supported by a grant from the Danish Natural Science Research Council, and by the C2VR of the Ecole Polytechnique at Palaiseau (France). W.D. thanks the Danish Space Board for travel support. NCAR is sponsored by the NSF.

Received 20 September; accepted 17 October 1988.

- Hummer, D. G. & Mihalas, D. *Astrophys. J.* **331**, 794–814 (1988).
- Mihalas, D., Däppen, W. & Hummer, D. G. *Astrophys. J.* **331**, 815–825 (1988).
- Däppen, W., Mihalas, D., Hummer, D. G. & Mihalas, B. W. *Astrophys. J.* **332**, 261–270 (1988).
- Seaton, M. J. *Phys. B20*, 6363–6378 (1987).
- Ebeling, W., Kraeft, W. D., Kremp, D. & Röpke, G. *Astrophys. J.* **290**, 24–27 (1985).
- Rogers, F. *Astrophys. J.* **310**, 723–728 (1986).
- Däppen, W., Anderson, L. S. & Mihalas, D. *Astrophys. J.* **319**, 195–206 (1987).
- Christensen-Dalsgaard, J., Duvall, T. L., Gough, D. O., Harvey, J. W. & Rhodes, E. J. *Nature* **315**, 378–382 (1985).
- Deubner, F.-L. & Gough, D. O. *A. Rev. Astr. Astrophys.* **22**, 593–619 (1984).
- Christensen-Dalsgaard, J. *Seismology of the Sun and the Distant Stars* (ed. Gough, D. O.) 23–53 (Reidel, Dordrecht, 1986).
- Christensen-Dalsgaard, J. *Multimode Stellar Pulsation* (eds Kovacs, G., Szabados, L. & Szeidl, B.) 153–180 (Kultura Press and Konkoly Observatory, Budapest, 1988).
- Grec, G., Fossat, E. & Pomerantz, M. A. *Sol. Phys.* **82**, 55–66 (1983).
- Libbrecht, K. G. *Astrophys. J.* **334**, 510–516 (1988).
- Ando, H. & Osaki, Y. *Publ. astr. Soc. Japan* **27**, 581–603 (1975).
- Christensen-Dalsgaard, J., Gough, D. O. & Libbrecht, K. G. *Astrophys. J.* (submitted).
- Christensen-Dalsgaard, J. *Méthodes mathématiques pour l'astrophysique* (eds Auvergne, M. & Baglin, A.) 479–524 (Société Française des Spécialiste d'Astronomie, Paris, 1987).
- Eggleton, P. P., Faulkner, J. & Flannery, B. P. *Astr. Astrophys.* **23**, 325–330 (1973).
- Christensen-Dalsgaard, J. *Mon. Not. R. astr. Soc.* **199**, 735–761 (1982).
- Lebreton, Y. & Däppen, W. *Proc. Symp. on Seismology of the Sun and Sun-like Stars* (ESA SP-286, ESTEC, Noordwijk, in the press).
- Christensen-Dalsgaard, J., Gough, D. O. & Pérez Hernández, F. *Mon. Not. R. astr. Soc.* (in the press).
- Duvall, T. L., Harvey, J. W., Libbrecht, K. G., Popp, B. D. & Pomerantz, M. A. *Astrophys. J.* **324**, 1158–1171 (1988).
- Libbrecht, K. G. & Kaufman, J. M. *Astrophys. J.* **324**, 1172–1183 (1988).
- Christensen-Dalsgaard, J. & Gough, D. O. *Solar Seismology from Space* (eds Ulrich, R. K. *et al.*) 199–204 (NASA, JPL. Publ. 84-84, 1984).
- Noels, A., Scuflaire, R. & Gabriel, M. *Astr. Astrophys.* **130**, 389–396 (1984).
- Lebreton, Y. & Maeder, A. *Astr. Astrophys.* **161**, 119–124 (1986).
- Bahcall, J. N. & Ulrich, R. K. *Rev. mod. Phys.* **60**, 297–372 (1988).
- Turck-Chieze, S., Cahen, S., Casse, M. & Doom, C. *Astrophys. J.* **335**, 415–424 (1988).
- Däppen, W. & Gough, D. O. *Theoretical Problems in Stellar Stability and Oscillations* 264–269 (Institut d'Astrophysique, Liège, 1984).
- Christensen-Dalsgaard, J. & Frandsen, S. *Sol. Phys.* **82**, 165–204 (1983).
- Gough, D. O. *Adv. Sp. Res.* **4**, 85–102 (1985).
- Gough, D. O. *Future Missions in Solar, Heliospheric and Space Plasma Physics* (eds Rolfe, E. & Battick, B.) 183–197 (ESA SP-235, ESTEC, Noordwijk, 1985).
- Stein, R. F., Nordlund, Å. & Kuhn, J. R. *Solar and Stellar Granulation Proc. NATO ARW, Capri*, 21–25 June (in the press).

A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons

Chand Desai, Gian Garriga, Steven L. McIntire* & H. Robert Horvitz

Howard Hughes Medical Institute, Department of Biology, Room 56-623, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

* Program in Neuroscience, Harvard Medical School, Boston, Massachusetts 02115, USA

Thirty-five genes define a pathway for the development of the hermaphrodite-specific neurons (HSNs) in Caenorhabditis elegans. Some of these genes affect only one HSN trait, demonstrating that HSN migration, axonal outgrowth and serotonin expression are mutually independent events in HSN development; others, some of which are regulatory, affect multiple HSN traits. Nearly all are pleiotropic, revealing that the genes specifying HSN development also function in the development of other cell types.

How are different cell types specified during development? To address this question, we are analysing the genes required for the determination, differentiation and function of the serotonergic HSNs of the nematode *C. elegans*. The HSN motor neurons are known to be necessary for egg-laying by *C. elegans* hermaphrodites, as elimination of these cells either by mutation or with a laser microbeam causes animals to be egg-laying defective (Egl)^{1–3}. Here, we describe the development of the HSNs, the identification and characterization of HSN-defective mutants, and a genetic pathway for HSN development.

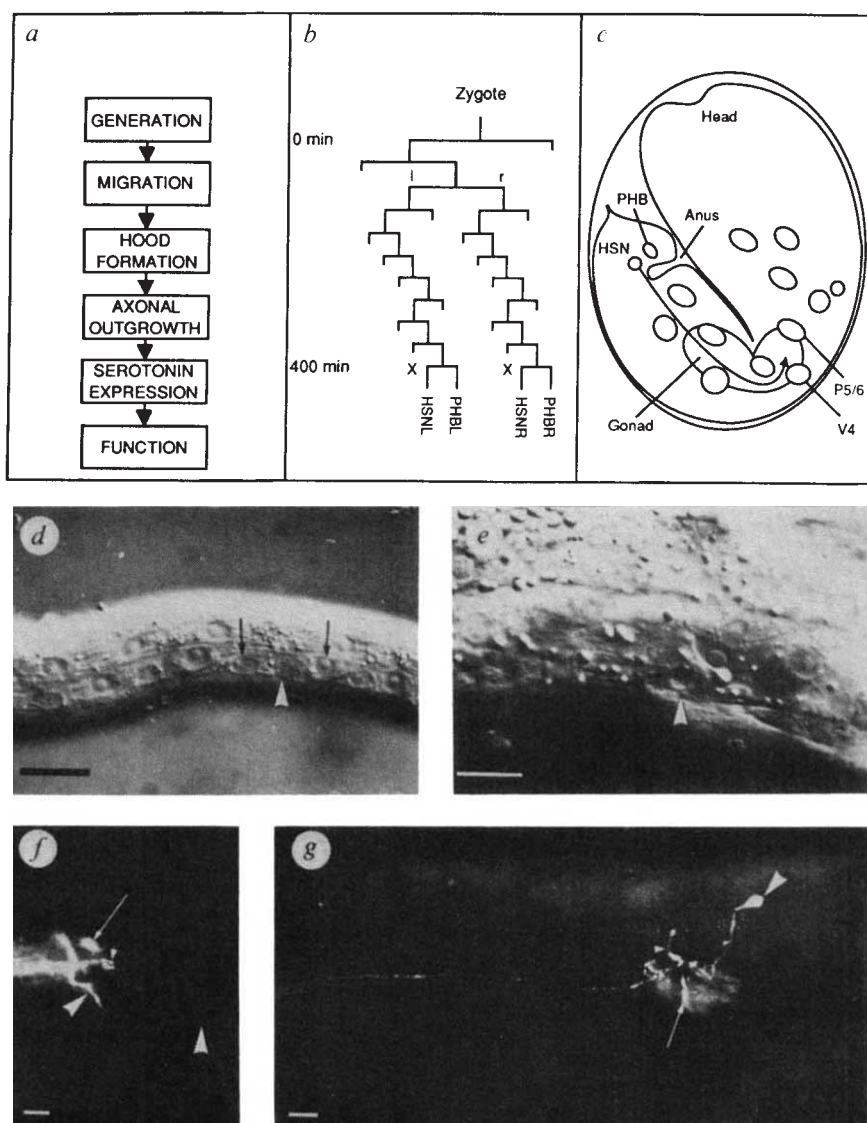
Wild-type HSN development

The major events of HSN development are illustrated in Fig. 1. The two bilaterally symmetric HSNs are generated halfway through embryogenesis, eleven cell divisions after fertilization (Fig. 1b). Ten minutes later the HSNs begin to migrate anteriorly

from their birthplace in the tail. In the hermaphrodite, the HSNs migrate to positions about midway along the length of the embryo (Fig. 1c, d). In the male, the homologous cells undergo programmed cell death during this migration, about an hour after their generation.

As visualized by Nomarski microscopy, the HSNs look like typical *C. elegans* neurons during the first three larval stages (L1–L3): their nuclei are small, round, stippled and refractile (Fig. 1d). During the fourth larval stage (L4), the HSNs undergo morphological maturation: the HSN nucleus and nucleolus enlarge and a distinctive eye-shaped structure, called the 'hood', forms around the nucleus (Fig. 1e). HSN axonal outgrowth is incomplete at this stage, on the basis of the analysis of electron micrographs of serial sections of an L4 hermaphrodite⁴. By the young adult stage, HSN axonal outgrowth is complete, as immunocytochemical staining with antisera to serotonin shows

Fig. 1 Wild-type development of the HSN neurons. *a*, A temporal pathway of wild-type HSN development. All developmental stages shown can be observed by light microscopy. Generation, migration and hood formation can be seen using Nomarski optics. Axonal outgrowth (indicated by the morphology of the HSN axon) and serotonin expression (indicated by the presence of immunocytochemically-detectable serotonin) can be observed in fixed animals after staining with antisera to serotonin. Function is inferred from the ability of animals to lay eggs. *b*, The embryonic cell lineage that gives rise to HSNL and HSNR, the left and right HSNs, respectively. The vertical axis indicates time at 20° from the first cleavage of the fertilized egg (0 min). The HSN arises at 400 min. One cell division shown is along the left-right axis; the daughters are labelled 'l' and 'r'; all other divisions are along the anterior-posterior axis, with cells shown to the left being anterior to cells shown to the right. X, programmed cell death, which occurs shortly after cell division. The HSN sister PHB is a sensory neuron in the phasmid. *c*, Schematic lateral view of an embryo, 400 min after fertilization, at the time of HSN birth. Only a few of the cells (indicated by small ovals and circles) present at this time are shown. P5/6 and V4 are blast cells that divide postembryonically³⁵. The final position of the HSN is between P5/6 and V4. The arrow indicates the route of HSN migration; however, when the HSN arrives at its final destination, the animal is much longer than shown here. *b, c* After ref. 36. *d-g*, Anterior is to the left; ventral is down. *d*, Nomarski photomicrograph of the left side of a recently hatched L1 stage wild-type (N2) hermaphrodite. The HSNL nucleus is marked by an arrowhead. The nucleus immediately anterior to that of the HSN belongs to P5/6L, whereas the nucleus immediately posterior belongs to V4L (marked by arrows). *e*, Nomarski photomicrograph of the left side of an L4 stage N2 hermaphrodite. The HSNL nucleus is marked by an arrowhead. Note the distinctive eye-shaped hood surrounding the nucleus. *f*, A fluorescence photomicrograph of the left side of the head of a fixed adult *unc-31(e169)* hermaphrodite, after staining with an antiserum to serotonin. Serotonergic neurons of *unc-31* animals are morphologically indistinguishable from those of the wild-type but appear to stain more brightly. The HSN axon (large arrowheads) runs along the ventral side of the animal and then turns and enters the nerve ring. The staining oval located posterior to the HSN axon in the nerve ring (arrow) is the cell body of a sensory neuron in the amphid. The processes marked by the small arrowheads belong to the NSM neurons, which are located in the pharynx^{4,37}. *g*, A fluorescence photomicrograph of the left side of a fixed adult N2 hermaphrodite, after staining with an antiserum to serotonin. The bright oval (large arrowhead) at the right is the HSNL cell body. An axon emerges from the anterior tip of the cell body and then dips ventrally. After entering the left side of the ventral cord, the HSNL axon skirts the vulval opening (the bright vertical line marked by an arrow). At the peak of the vulval detour, the HSNL axon forms varicosities and a branch (small arrowhead). Anterior to the vulva, the HSNL axon rejoins the left side of the ventral cord and runs anteriorly. The HSNR axon can also be seen anterior to the vulva running parallel to the HSNL axon. Methods are as in the Fig. 2 legend. Scale bars, 10 µm.



(see below). A single axon emerges from each HSN cell body and extends ventrally into the ventral nerve cord (Fig. 1g), which consists of two parallel axonal bundles that extend the length of the animal⁴. The axons from the left and right HSNs enter the left and right bundles, respectively. Each HSN axon then turns anteriorly and runs past the vulva, where it forms varicosities (swellings) and usually branches (Fig. 1g). Many synapses from the HSNs to the egg-laying muscles occur on this branch⁴. The HSN axon continues anteriorly into the head, where it enters the nerve ring (Fig. 1f, g). In this dense neuropil, which encircles the pharynx, the HSN axons synapse with several classes of neurons⁴.

Immunocytochemical and pharmacological experiments indicate that serotonin is an HSN neurotransmitter. Indirect immunofluorescence microscopy reveals the presence of serotonin-like immunoreactive material in the HSNs (Fig. 1f, g).

Serotonin is first detectable in the HSNs of young adults. Both serotonin and an inhibitor of serotonin uptake, imipramine⁵, stimulate wild-type hermaphrodites to lay eggs². By contrast, animals lacking HSNs lay eggs in response to serotonin but not in response to imipramine^{2,3}. These observations suggest that serotonin stimulates the egg-laying muscles directly and that imipramine stimulates egg laying by potentiating HSN-derived serotonin (and for this reason has no effect on egg laying in the absence of HSNs). Furthermore, serotonin receptor antagonists inhibit egg laying, which indicates that endogenous serotonin is necessary for normal egg laying (E. Wolinsky, personal communication; C. Johnson, personal communication).

Identification of HSN-defective mutants

To identify genes responsible for the development of the HSN neurons, we used two approaches. Firstly, we screened for

Table 1 HSN-defective mutants

Genes	LG	No.	Notes	HSN defects	Other defects	Source
<i>cat-1</i>	X	1		5HT	Amines, Mate*	28†
<i>cat-4</i>	V	1		5HT	Amines, Mate*	28†
<i>egl-1</i>	V	4	ts, dm	Prog. death	Egl	2, 3
<i>egl-5</i>	III	7		Mig, Mat, 5HT	Egl, Mate, Mec(t), Unc	2, 3, d, e
<i>egl-10</i>	V	9	ts, sd		Egl, Unc, Mate*	2, 3
<i>egl-18</i>	IV	3		Mig	Dye, Egl, Mig*, Vulva	3†
<i>egl-20</i>	IV	2	ts*	Mig	Egl, Mate*, Unc	3, a†
<i>egl-27</i>	II	1		Mig	Dye, Egl, Mate, Mig	3†
<i>egl-41</i>	V	4	ts, dm	Prog. death	Eg., Sex	2, 10
<i>egl-42</i>	II	2	sd		Egl	3
<i>egl-43</i>	II	2		Mig, 5HT	Dye, Egl, Mate*	3
<i>egl-44</i>	II	3		Mig, Axon, 5HT	Egl, Mate*	3
<i>egl-45</i>	III	1		Mat, Axon, 5HT	Egl, Mate	3
<i>egl-46</i>	V	4		Mig, Axon, 5HT	Egl, Lin, Mate*, Unc	3
<i>egl-47</i>	V	2	dm		Egl, Mate*	3
<i>egl-49</i>	X	1			Egl	3
<i>egl-50</i>	II	1			Egl, Mate	3
<i>egl(n1108)</i>	V	1	ts		Egl, Mate*	3
<i>egl(n1653)</i>	III	1		Prog. Death	Egl, Sex	b
<i>ham-1</i>	IV	3		Mig, 5HT	Dye, Egl	c, e†
<i>ham-2</i>	X	1		Mig, 5HT	Egl	a
<i>ham-3</i>	III	1		Mig, 5HT	Egl, Mig*	e†
<i>her-1</i>	V	1	dm	Prog. death	Egl, Sex	2
<i>mig-1</i>	I	4		Mig	Egl*	11, a, e
<i>mig-2</i>	X	1		Mig, Axon	Egl, Mig, Unc	11†
<i>mig-10</i>	III	1		Mig	Dye, Egl, Mig, Vulva	f†
<i>mig-12</i>	IV	1		Mig	Egl, Mig, Unc, Vulva	g†
<i>sem-4</i>	I	1		Mat, Axon, 5HT	Egl, Lin	h†
<i>tra-2</i>	II	>20 (3)		Prog. death	Egl, Sex	3, 8
<i>unc-6</i>	X	4 (3)		Axon	Axon, Egl, Mate, Unc	12†
<i>unc-33</i>	IV	5 (1)		Axon	Axon, Egl, Mate, Unc	12†
<i>unc-34</i>	V	3 (2)		Axon	Axon, Egl-C, Mate, Unc	12†
<i>unc-40</i>	I	4 (3)		Axon	Axon, Egl, Mate, Unc	12†
<i>unc-51</i>	V	6 (2)		Axon	Axon, Egl, Mate, Unc	12†
<i>unc-71</i>	III	1		Mig, Axon	Axon, Egl*, Mate*, Unc	12†
<i>unc-73</i>	I	1		Mig, Axon	Axon, Egl, Mate, Unc	12†
<i>unc-76</i>	V	2		Axon	Axon, Egl-C, Mate, Unc	12, i†
<i>unc-86</i>	III	36 (9)		Mat, Axon, 5HT	Egl, Lin, Mec	3, 27, 29

Thirty-eight genes are required for HSN development and/or function. LG, linkage group to which the gene maps. No., number of mutant alleles of the gene. Numbers in parentheses refer to the number of different mutant strains examined for their HSN phenotypes. Unless indicated in parentheses, mutants representing all alleles were examined. Notes, ts, temperature-sensitive; ts*, weakly temperature-sensitive; dm, dominant; sd, semidominant. HSN defects: Prog. death, HSNs undergo embryonic programmed cell death in hermaphrodites; Mig, migration abnormal; Mat, L4 stage morphological maturation (hood formation and nucleolar growth) defective; Axon, axonal outgrowth abnormal; 5HT, HSNs lack or have reduced levels of serotonin by immunocytochemical staining; mutants without a listed HSN defect are thought to be HSN-defective because they are Egl and lay eggs in serotonin but not in imipramine. Other defects: Amines, reduced levels of serotonin and dopamine in all serotonergic and dopaminergic cells (ref. 30, and data not shown); Axon, defects in axonal outgrowth of cells other than the HSNs (refs 11, 13 and S.M. and H.R.H., unpublished observations); Dye, altered FITC loading of amphid and/or phasmid cells^{11,13}; Egl, the Egl phenotype can be caused either by HSN abnormalities or by abnormalities in other components of the egg-laying system³¹; Egl*, low penetrance egg-laying defective (4–20%); Egl-C, egg-laying constitutive (see text); Lin, cell-lineage defective; Mate, males do not mate³²; Mate*, reduced efficiency of male mating^{2,3,32}; Mec, mechanosensory defective^{20,33}; Mec(t), mechanosensory defective in the tail; Mig, defects in several cell migrations in addition to the HSN migration; Mig*, weakly penetrant defects in cell migration(s) in addition to the HSN migration; Sex, sexual fates of multiple sexually dimorphic cells altered^{2,3,7,8}; Unc, uncoordinated movement^{2,3,11,12}; Vulva, vulval abnormalities². Many mutants were isolated on the basis of a serotonin-sensitive imipramine-resistant Egl phenotype^{2,3} (also see text); some HSN-abnormal mutants were obtained using other methods. †, Mutant strains isolated on the basis of defects other than HSN abnormalities that were later shown to be HSN-abnormal by Nomarski or fluorescence microscopy. a, From 200,000 *mut-2*³⁴ hermaphrodites, 288 true-breeding Egl strains were identified that lay eggs in 7.5 mg ml⁻¹ serotonin; using Nomarski microscopy, we found abnormal HSNs in *egl-20(n1437)*, *ham-2(n1332)*, and *mig-1(n1354, n1652)* animals. b, *egl(n1653)* was isolated using Nomarski microscopy to screen 2,793 unselected L1-stage F2 progeny of ethyl methanesulphonate-mutagenized hermaphrodites. Of 52 animals isolated with missing or misplaced HSNs, only *egl(n1653)* survived and bred true. c, *n1810* and *n1811* were isolated after ethyl methanesulphonate mutagenesis on the basis of their failure to complement *ham-1(n1438)* for its Egl phenotype. Several of the mutants described in this table were isolated in screens not designed to identify HSN-abnormal mutants *per se*: d, *egl-5(u202)*, mechanosensory defective in the tail (M. Chalfie, personal communication); e, *egl-5(n1439)*, *ham-1(n1438)* (M. Stern, personal communication), *ham-3(n1654)* (M. Herman, personal communication), and *mig-1(n687)* (C. Trent, personal communication); Egl; f, *mig-10(ct41)* (J. Manser, personal communication), several embryonic cell migration defects; g, *mig-12(n1706)*, vulval protrusion (S. Federhen, personal communication); h, *sem-4(n1378)*, transformation of sex myoblasts into body wall muscles (M. Stern, personal communication); i, *unc-76(ev424)*, Unc (S. Siddiqi, personal communication).

mutants defective in HSN function^{2,3}. Hermaphrodites that lack HSNs are Egl and are stimulated to lay eggs by serotonin but not by imipramine. By contrast, hermaphrodites that lack the VC neurons, the only neurons besides the HSNs to innervate the egg-laying muscles, are not Egl; hermaphrodites that lack the egg-laying muscles or vulvae are Egl but do not lay eggs in

response to serotonin or to imipramine. In a screen for animals impaired in HSN function, we isolated 47 serotonin-sensitive, imipramine-resistant Egl mutants. These mutations define 16 genes involved in HSN development^{2,3}.

The second approach involved screening for mutants with morphologically abnormal or missing HSNs. Mutants represent-

Fig. 2 Mutant HSN phenotypes. Anterior is to the left. *a, b*, Nomarski photomicrographs of living animals. *c–g*, Fluorescence photomicrographs of animals stained with antisera to serotonin. *a*, Migration-defective, young L1 *egl-20(n1437)* hermaphrodite (left lateral view). The arrowhead marks the posteriorly displaced HSNL nucleus. The arrows mark, from top to bottom, the nuclei of CANL neuron and V4L, which flank the HSNL nucleus in the wild-type hermaphrodite (the CANL neuron lies just dorsal to P5/6L, which is slightly out of the plane of focus in this photomicrograph; compare with Fig. 1*d*). *b*, Maturation-defective, L4 *egl-45(n999)* hermaphrodite (left lateral view). The arrowhead marks the nucleus of a degenerating HSNL. Note the absence of a hood and the overall faint appearance of this HSN nucleus and nucleolus (compare with Fig. 1*e*). *c*, Axonal outgrowth-defective, *unc-51* adult (ventral view). The HSN cell bodies are marked by arrowheads. The HSNR axon forms several large varicosities, marked by arrows. Neither the HSNL nor HSNR axon runs anteriorly in the ventral cord. *d*, Axonal outgrowth-defective, *unc-40(n324)* adult (left lateral view). The HSNL axon (small arrowheads) exiting from the HSN cell body (large arrowhead) does not enter the ventral cord, and instead projects anteriorly along the lateral hypodermis. *e*, Axonal outgrowth-defective, *unc-34(e566)* adult (left lateral view). The HSNL axon (marked by arrowheads) enters the ventral cord and runs anteriorly but does not reach the nerve ring. *f*, Identity-defective, *ham-1(n1810)* adult (left lateral view). The four intensely staining circles marked by arrowheads are HSN-like cell bodies. All four are migration-defective; three are located in the tail. The vulva is located anterior to the anterior-most region shown in this photomicrograph. The apparent longitudinal staining is of the lumen of the intestine, which occasionally displays fluorescence in this immunocytochemical staining protocol; this fluorescence is not specific to anti-serotonin antisera. *g*, *egl-43(n997)* adult in which the HSNs stain with an antiserum to serotonin (ventral view). The brightly staining ovals marked by arrowheads near the tail are HSN cell bodies that have migrated about 35% of the distance from the anus to the vulva. The vulva is the transverse line marked by an arrow. The HSN axons continued to the nerve ring in this animal. Two axonal outgrowth defects are present, which are generally seen in posteriorly displaced HSNs. First, after both HSN axons enter the ventral cord, the HSNL axon crosses the hypodermal ridge and runs on the right side of the ventral cord with the HSNR axon (compare with Fig. 1*g*). Second, the HSN axons fail to form varicosities and fail to branch at the vulva (compare with Fig. 1*g*). The striped pattern is caused by non-specific staining between eggs. Scale bars, 10 μ m.

Methods. Animals were prepared for Nomarski microscopy and photographed as described³⁵. Indirect immunofluorescence histochemistry was used to stain animals for serotonin. Adult hermaphrodites were fixed in 4% paraformaldehyde fixative for 24 h at 4°. After washing in PBS (50 mM Na₂HPO₄, 140 mM NaCl, pH 7.2), the worms were rocked gently overnight at 37° in a solution of 5% β -mercaptoethanol, 1% Triton X-100 in 0.125 M Tris HCl pH 6.9. The worms were washed in PBS and 20–30 μ l of worms were shaken vigorously in 0.4 ml of 100 mM Tris HCl pH 7.5, 1 mM CaCl₂, and 1,000–2,000 units ml⁻¹ collagenase (Sigma type IV). After 20–50% of the worms were broken (1–3 h), the worms were washed with PBS and stained using an anti-serotonin primary antiserum and FITC-conjugated secondary and tertiary antibodies. 5 μ l of stained worms were mixed with 5 μ l of 1 mg ml⁻¹ *p*-phenylenediamine in 10% PBS pH 8.0 and 90% glycerol and viewed by fluorescence microscopy (illuminated with a mercury vapour lamp, and viewed through an FITC filter; Zeiss 48705, excitation range 395–440 Å).

ing 11 *egl* (egg-laying defective), *ham* (HSN abnormal migration) and *mig* (migration-defective) genes were identified as having abnormal HSNs by examining live animals using Nomarski microscopy. Existing *cat* (catecholamine abnormal), *sem* (sex muscle abnormal) and *unc* (uncoordinated) mutants representing 11 genes were identified as HSN-abnormal by indirect immunofluorescence microscopy of fixed animals stained with antisera to serotonin. Additional alleles of many of these genes were isolated in unrelated experiments. Altogether, mutations in 38 genes have been shown to affect the HSNs (Table 1). Characterization of these mutations reveals that they disrupt most steps in HSN development.

Steps affected in HSN-defective mutants

(1) **Sex-specific HSN survival.** In our mutant screens, we identified alleles of *egl-41*, *her-1*, *tra-2*, *egl(n1653)* and *egl-1* that affect HSN sex-specific survival (*her*, transformer of XO animals—normally males—into hermaphrodites; *tra*, transformer of XX animals—normally hermaphrodites—into males). Hermaphrodites mutant in these five genes often hatch without HSNs. The HSNs are restored in double-mutant combinations between these genes and *ced-3(n717)*. As *ced-3* mutations prevent the onset of programmed cell death⁶, the survival of the HSNs in these double mutants indicates that the HSNs undergo

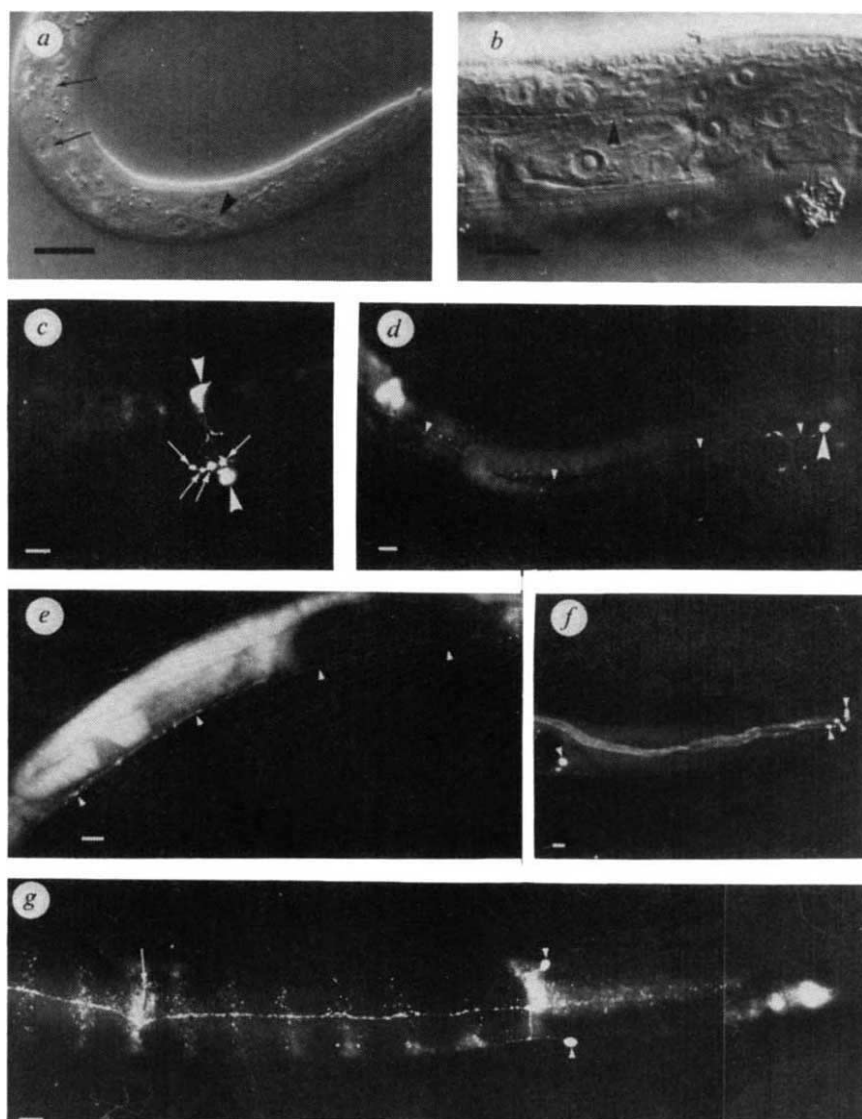


Table 2 The effects of mutations on HSN migration and egg laying

HSN-migration mutants			HSN-migration mutants with other HSN defects		
Genotype	Position of HSN cell body (n)	% Egl (n)	Genotype	Position of HSN cell body (n)	% Egl (n)
wild-type	100 ± 1% (48)	<1 (>1,000)	<i>egl-43</i> (n997)	5 ± 2% (158)	90 (238)
<i>egl-20</i> (n1437)	70 ± 4% (141)	48 (606)	(n1079)	10 ± 3% (136)	92 (242)
(n585)	72 ± 1% (170)	89 (398)	<i>egl-5</i> (u202)	24 ± 5% (102)	100 (896)
<i>mig-10</i> (ct41)	70 ± 4% (143)	78 (263)	(n1066)	25 ± 5% (99)	100 (804)
<i>mig-1</i> (n1354)	69 ± 5% (71)	7 (661)	(n945)	28 ± 3% (299)	100 (918)
(n1652)	76 ± 3% (94)	2 (1,023)	(n988)	32 ± 4% (121)	100 (960)
(e1787)	78 ± 2% (185)	20 (878)	(n989)	38 ± 4% (173)	98 (912)
<i>egl-27</i> (n170)	77 ± 4% (143)	80 (113)	(n486)	46 ± 5% (117)	97 (888)
<i>egl-18</i> (n474)	73 ± 5% (134)	78 (154)	(n1439)	52 ± 5% (119)	92 (1,074)
(n475)	74 ± 5% (137)	77 (183)	<i>ham-1</i> (n1438)	41 ± 9% (84)	57 (451)
(n162)	87 ± 4% (139)	14 (423)	(n1810)	41 ± 7% (134)	40 (1,475)
<i>mig-12</i> (n1706)	†	29 (515)	(n1811)	51 ± 7% (131)	39 (1,300)
			<i>ham-3</i> (n1654)	54 ± 5% (105)	88 (270)
			<i>ham-2</i> (n1332)	59 ± 3% (148)	88 (476)
			<i>mig-2</i> (rh17)	76 ± 2% (117)	97 (372)
			<i>unc-73</i> (e936)	84 ± 3% (94)	34 (1,125)
			<i>unc-71</i> (e541)	86 ± 2% (81)	4 (721)

% Egl, the percentage of adults that retain eggs as scored with a dissecting microscope. Position of HSN cell body, average position of the HSN ± s.e.m., scored using Nomarski microscopy by marking the location of the HSN nucleus on a cartoon of an L1 hermaphrodite divided into 40 regions between the anus (0.00 migration) and the middle of the gonadal primordium (1.00 migration). All numbers were then normalized to the average HSN position in wild-type animals (0.92), so that the wild-type value is defined as 100%. Those HSNs not in this interval were assumed to be present in the lumbar ganglia posterior to the anus, and were scored as having migrated 0%. In mutants in which the HSN does not contain detectable serotonin (for example *egl-5*), however, we could not confirm the presence of postanal HSNs by immunofluorescence; in such mutants, the HSNs may not have been generated or may have died and degenerated. † In *mig-12* L1 animals it is often impossible to score the HSN positions relative to marker cells, because many of the marker cells are mispositioned, but the HSN migration defect can be seen clearly in *mig-12* adults stained with antisera to serotonin; we estimate that the position of the HSNs in *mig-12* animals is about 70%. *egl-44* and *egl-46* animals, which have subtle abnormalities in HSN position, were scored during the L4 stage, by a process that allows a more sensitive measure of HSN position. L4 HSN position was determined using Nomarski microscopy by scoring the position of the anterior edge of the HSN hood (see Fig. 1e) relative to the vulva; positions anterior to the vulva have positive values, positions posterior to the vulva have negative values, and the position of the vulva is scored as 0.0. Units are average HSN cell diameters (approximately 10 µm). The following data indicate the average HSN position ± s.e.m.: N2: HSNR = -1.0 ± 0.1, HSNL = -2.1 ± 0.2; *egl-44*: HSNR = -0.4 ± 0.1, HSNL = -1.6 ± 0.2; *egl-46*: HSNR = -0.5 ± 0.2, HSNL = -1.3 ± 0.2. In all three strains, the right HSN is located slightly anterior to the left HSN; a left-right HSN difference of this magnitude would not be detectable in L1 animals scored as described above. The anterior shift of HSN positions in *egl-44* and *egl-46* mutants can be detected more obviously by examining the positions of HSNs in L1 double mutant animals also defective in *egl-43*: *egl-43*(n997): 5 ± 2%; *egl-44*(n1080) *egl-43*(n997): 24 ± 7%; *egl-43*(n997); *egl-46*(n1127): 17 ± 6%. As these effects can be seen in L1 animals, they presumably result from a defect in embryonic HSN migration.

embryonic programmed cell death in mutants defective in these five genes^{2,3,7}.

As mentioned above, embryonic programmed cell death is the normal fate of the HSNs in males. The expression of the male HSN fate in these mutant hermaphrodites suggested that these five genes are involved in determining the sexual fate of the HSNs^{2,3}. Indeed, *her-1* and *tra-2* are major sex-determining genes of *C. elegans*^{8,9}, and the *her-1* and *tra-2* mutations that affect sex-specific HSN survival seem to be unusual alleles of these genes^{2,3,7}. Mutations in *egl-41* and *egl-1653* result in apparent intersexual animals, and these genes may also function in the general sex-determination pathway^{3,10}. In contrast to these four genes, *egl-1* seems to affect only the HSNs and may mediate the interaction between the sex-determination pathway and HSN development^{2,3}.

(2) **HSN migration.** Mutations in six genes disrupt HSN migration without affecting other visible HSN traits. The HSN cell bodies in these mutants are located at various points along the HSN migratory route posterior to their position in wild-type animals (Fig. 2a and Table 2, left). Even when located in the postanal region of the tail, the HSNs express serotonin and extend axons into the ventral cord and then anteriorly into the nerve ring. *mig-10* and *mig-12* mutations affect a number of other migrations, including those of the ALM and CAN neurons (the other two neuron types with long migrations) (ref. 11, J. Manser, personal communication). In addition, in *mig-12* animals the positions of many other cells are abnormal (G. G. and M. Herman, unpublished observations). These two genes may define components involved in the mechanics of cell migra-

tion. By contrast, *egl-18*, *egl-20*, *egl-27* and *mig-1* mutations affect primarily HSN migration.

Among the different HSN-migration mutants, the severity of the migration defect does not correlate with the penetrance of the egg-laying defect (Table 2, left). For example, the HSNs in both *mig-1*(n1652) and *egl-27* mutants migrate about 77% as far as the HSNs in wild-type animals; however, only 2% of *mig-1*(n1652) hermaphrodites are Egl, whereas 80% of *egl-27* animals are Egl. This lack of correlation indicates that severely Egl HSN-migration mutants have additional defects. *egl-18*, *egl-20*, *egl-27*, *mig-10* and *mig-12* mutants probably have defects in the egg-laying muscles and/or vulva, as they frequently fail to lay eggs in serotonin (ref. 2 and data not shown). These observations suggest that the HSNs may be functionally normal in some, and perhaps all, of the HSN migration mutants.

(3) **Hood formation.** Mutations in three genes disrupt the morphological maturation that the HSNs undergo during the L4 stage. At hatching, the HSNs in *egl-45*, *sem-4* and *unc-86* mutants appear normal by Nomarski microscopy. During the L4 stage, the HSNs in these mutants generally fail in nucleolar growth and hood formation (Fig. 2b); these HSNs often become indistinct or undetectable. In addition, HSN serotonin is not detectable in *unc-86* adults and is only occasionally detectable in *egl-45* and *sem-4* adults. In those *egl-45* and *sem-4* individuals that have detectable HSN serotonin, the HSNs often extend aberrant axons. The phenotypes of these mutants suggest that they are defective either in a regulatory step required for all late aspects of HSN differentiation—nucleolar growth, hood formation, axonal outgrowth, serotonin expression, and function—or

in hood formation *per se*, which in this case must be necessary for the expression of later-appearing HSN traits. The former hypothesis is strongly supported by other studies of *sem-4* and *unc-86* (see conclusions).

(4) **Axonal outgrowth.** Mutations in six genes affect HSN axonal outgrowth without altering other visible HSN traits (Table 3, top). Mutations in four of these genes specifically disrupt either ventrally-directed HSN axonal growth to the ventral cord or anteriorly-directed HSN axonal growth within the ventral cord. Mutations in two other genes perturb several aspects of HSN axonal outgrowth or morphology. Mutations in all six of these genes cause uncoordinated locomotion¹², which probably results from axonal-outgrowth defects of other neurons (refs 11, 13; and S.M. and H.R.H., unpublished observations).

unc-6 and *unc-40* mutants are defective in the growth of the HSN axons to the ventral cord. In these mutants, the HSN axons generally project anteriorly along a novel lateral pathway to reach the nerve ring, their normal destination (Fig. 2d). The *unc-6* gene has been molecularly cloned, and the sequence of the putative *unc-6* protein is similar to that of the B2 chain of laminin, suggesting that the *unc-6* protein functions directly in axonal guidance (N. Ishii, B. Stern, J. Culotti and E. Hedgecock, personal communication).

The HSN axons in *unc-34* and *unc-76* mutants grow ventrally and form varicosities and a branch at the vulva, as in the wild-type. These axons are defective in anterior growth however and terminate at variable points along the ventral cord (Fig. 2e). Interestingly, neither *unc-34* nor *unc-76* animals are blocked in egg laying. Instead, unlike wild-type animals, which retain eggs when removed from bacteria¹⁴, *unc-34* and *unc-76* mutants lay eggs even in the absence of bacteria (E. Wolinsky, personal communication; C.D., G.G., S.M. and L. Bloom, unpublished observations). This constitutive phenotype for egg laying may result from the loss of inhibitory synaptic input to the anterior region of the HSN axon, perhaps in the nerve ring. Such input might modulate the rate of egg laying in response to bacteria in the environment¹⁴.

Like *unc-34* and *unc-76* animals, *unc-51* animals have HSN axons defective in anterior growth. However, *unc-51* animals have a serotonin-sensitive and Egl phenotype, suggesting that their HSNs are impaired in their ability to stimulate egg-laying. *unc-51* mutants display additional visible HSN defects, including ectopic, abnormally large varicosities along their axons and occasional extra neurites extending from their cell bodies (Fig. 2c).

unc-33 mutants display less specific disruptions of HSN axonal outgrowth. The HSN axons of *unc-33* animals generally either grow anteriorly along the lateral hypodermis or stop prematurely within the ventral cord.

In addition to the mutations described above, mutations in many genes that affect the vulva^{15,16} result in relatively minor alterations in HSN axonal morphology. As the HSN axons grow along the vulval cells, these HSN defects are probably a consequence of the vulval abnormalities (G.G., C.D. and H.R.H., unpublished observations; C. Li and M. Chalfie, personal communication).

(5) **Serotonin expression.** Mutations in *cat-1* and *cat-4* decrease immunocytochemical staining of serotonin in the HSNs as well as in the other eight serotonergic cells, which are present in the head (data not shown). Surprisingly, *cat-1* and *cat-4* mutants are not obviously Egl. Serotonin is reduced but clearly present in *cat-1* mutants, which presumably accounts for their ability to lay eggs. *cat-4* mutants have no immunocytochemically-detectable serotonin, however. Nevertheless, two observations suggest that *cat-4* mutants probably also contain residual serotonin. First, *cat-4* mutants lay eggs in response to imipramine (C.D., unpublished data). Second, imipramine stimulates pharyngeal pumping in both wild-type and *cat-4* animals (L. Avery and H.R.H., manuscript in preparation). As we believe imipramine acts in *C. elegans* by potentiating endogenous

serotonin, these observations suggest that *cat-4* mutants do not completely lack serotonin. Egg laying by *cat-4* mutants is hypersensitive to serotonin, indicating that abnormally low levels of endogenous serotonin could be adequate to stimulate egg laying in this strain (C.D., unpublished observations).

(6) **HSN function.** *egl-10*, *egl-42*, *egl-47*, *egl-49*, *egl-50* and *egl(n1108)* mutants are Egl and lay eggs in response to serotonin but not imipramine, which strongly suggests that their HSNs are functionally defective. The HSNs in these mutants however appear normal by Nomarski and fluorescence microscopy. To examine HSN synapses, we have begun to analyse these mutants by serial section electron microscopy. So far, we have studied a single *egl-42(n995)* hermaphrodite for HSN ultrastructural defects in the region of the vulva; these HSNs seem normal¹⁷. The HSNs in mutants of this class may be defective in function but not in structure. Consistent with this hypothesis, *egl(n1108)* animals have a temperature-sensitive period for abnormal egg laying that is after the HSNs have completed development³.

(7) **Mutants with multiple HSN defects.** Ten genes are required both for HSN migration and for other aspects of HSN development. *egl-5*, *egl-43*, *egl-44*, *egl-46*, *ham-1*, *ham-2* and *ham-3* mutants are abnormal in HSN migration and serotonin expression (Tables 1 and 2, right). At least some of these genes also affect axonal outgrowth (see below). *mig-2*, *unc-71* and *unc-73* mutants are abnormal in HSN migration and axonal outgrowth (Tables 1 and 2, right and 3, bottom). As defects in HSN migration, axonal outgrowth and serotonin expression can occur separately, these three traits must result from mutually independent processes. Mutations in genes that affect more than one of these traits could define regulatory steps in HSN development. The ten genes of this type can be separated into three classes, on the basis of the HSN traits they affect and genetic interactions.

(i) The genes *ham-1* and *egl-5* seem to control the acquisition of HSN identity. *ham-1* may be involved in the segregation of developmental potential between the HSN and its sister cell, the phasmid neuron PHB (Fig. 1b). *ham-1* animals have from zero to four cells that are HSN-like in character: these cells migrate; extend axons into the ventral cord and anteriorly to the nerve ring; express serotonin in adults; and are affected by genes (*egl-1*, *egl-5*, *unc-86* and *egl-45*) that affect HSNs (Fig. 2f; also, data not shown). The extra HSN-like cells are probably the sisters of the HSNs. When wild-type animals are exposed to the fluorescent dye fluorescein isothiocyanate (FITC), the PHBs take up the dye and can be visualized by fluorescence microscopy¹³. In *ham-1* mutants, the PHBs are usually not detected by this technique (data not shown). Thus, in *ham-1* animals there are extra HSN-like cells, and the HSN sisters are abnormal or missing. These results suggest that PHB is variably transformed to express the fate of its sister cell, the HSN. On the other hand, the HSNs are also abnormal in *ham-1* mutants: they are migration-defective and often fail to stain for serotonin. Perhaps the HSN is abnormal because it is being variably transformed into a PHB cell type. If so, *ham-1* could be required for the segregation of developmental potential between the HSN and PHB cells.

In *egl-5* mutants, the presumptive HSNs seem to be defective in all aspects in HSN development. Migration, hood formation, serotonin expression, and function are clearly blocked. Because of the lack of serotonin, we were unable to score axonal morphology in *egl-5* animals. In addition, *egl-5* mutants are abnormal in HSN programmed cell death: we estimate that the presumptive HSNs survive about 40% of the time both in *egl-5(n945)* males and in *egl-1(n986)*; *egl-5(n945)* hermaphrodites (data not shown). These observations suggest that HSN development is often not initiated in *egl-5* animals.

(ii) The genes *egl-43*, *egl-44*, *egl-46*, *ham-2* and *ham-3* affect HSN migration and serotonin expression. At least *egl-44* and *egl-46* mutants are also abnormal in HSN axonal outgrowth (Table 3, bottom). We do not know if mutations in the other

Table 3 The effects of mutations on HSN axonal outgrowth

HSN axonal outgrowth mutants		HSN axonal growth defects			
Genotype	Ventral blocks	Anterior blocks	Posteriorly misdirected	No defects	No.
wild-type	0%	0%	0%	100%	>1,000
<i>unc-6</i> (<i>e7</i>)	90%	0%	6%	10%	49
(<i>e78</i>)	15%	0%	2%	85%	92
<i>unc-40</i> (<i>e271</i>)	59%	1%	8%	48%	113
(<i>n324</i>)	82%	0%	0%	18%	50
(<i>n473</i>)	81%	0%	0%	19%	48
<i>unc-34</i> (<i>e315</i>)	0%	57%	0%	43%	44
(<i>e566</i>)	0%	67%	0%	33%	49
<i>unc-76</i> (<i>e911</i>)	2%	98%	0%	0%	55
(<i>ev424</i>)	7%	92%	1%	0%	115
<i>unc-51</i> (<i>e369</i>)	1%	97%	3%	0%	78
(<i>e1120</i>)	0%	100%	0%	0%	39
<i>unc-33</i> (<i>e204</i>)	20%	62%	9%	4%	45
HSN axonal outgrowth mutants with other HSN defects					
<i>egl-44</i> [†] (<i>n1080</i>)	20%	0%	11%	70%	197
(<i>n1087</i>)	4%	0%	8%	80%	25
<i>egl-46</i> [†] (<i>n1075</i>)	11%	0%	5%	72%	64
(<i>n1076</i>)	6%	0%	9%	79%	33
(<i>n1126</i>)	3%	0%	14%	84%	37
(<i>n1127</i>)	14%	0%	3%	78%	37
<i>mig-2</i> (<i>rh17</i>)	45%	50%	7%	5%	42
<i>unc-71</i> (<i>e541</i>)	0%	8%	15%	77%	52
<i>unc-73</i> (<i>e736</i>)	51%	37%	6%	9%	35

HSN axons were scored using fluorescence microscopy after staining with antisera to serotonin (see Fig. 2 legend). Ventral blocks: axons failed to extend ventrally from their cell bodies to the ventral cord and instead extended in a sublateral or lateral position. Some of these axons grew anteriorly and eventually entered the ventral cord in the proximity of a circumferential bundle of nerve processes, the BDU commissure, just posterior to the pharynx. Anterior blocks: axons entered the ventral cord normally and initiated anterior growth, but terminated before entering the nerve ring. Posteriorly misdirected: axons grew posteriorly either along the lateral hypodermis or in the ventral cord and never reached the nerve ring. No defects: axons entered the ventral nerve cord and projected to the nerve ring. No., the number of HSN axons scored. HSN axons that grew posteriorly a short distance before entering the ventral cord and later projected to the nerve ring were not included in any of the above categories; relatively few HSNs (generally fewer than 1% and always fewer than 10%) were of this type. [†] Few *egl-44* and *egl-46* axons are visible after staining with antisera to serotonin; the morphology of non-staining axons is unknown.

three genes affect HSN axonal outgrowth, as we have examined only the axons of those HSNs that stain with antisera to serotonin; very few *ham-3* HSNs stain, and although a considerable fraction of *egl-43* and *ham-2* HSNs stain and appear normal, it is possible that the axons of non-staining HSNs are abnormal. Hood formation in these mutants is normal. Sex-specific survival is also normal: in double mutants between *egl-1* (which causes HSNs to die) and each of these genes, the HSNs die.

Each of these genes either could act separately in HSN migration, serotonin expression and axonal outgrowth (for example, by affecting cytoskeletal or membrane elements involved in these processes) or could act to coordinately regulate these processes. We believe that the second possibility is more likely, because of the relative specificity of these genes for the HSNs. In particular, the HSNs are the only cells in these mutants for which serotonin levels and/or migration are affected, and are the only cells for which axonal outgrowth defects have been observed. We view the mutants in this category as being defective in the coordinate control of the expression of a specific subset of HSN traits.

The migration defects of *egl-44* and *egl-46* mutants are distinctive, as the HSNs in these mutants migrate along their normal pathways but terminate slightly beyond their normal final positions (see Table 2 legend).

(iii) The genes *mig-2*, *unc-71* and *unc-73* affect HSN migration and axonal outgrowth. As discussed above, such multiple effects could exist either because each of these genes acts separately in HSN migration and axonal outgrowth or because each acts to coordinately control these two processes. Two observations favour the former hypothesis. First, the migration and/or

axonal outgrowth of many cell types are abnormal in mutants defective in these three genes¹¹. Second, cell migration and axonal outgrowth are phenomenologically and mechanistically similar, and are known to require similar cytoskeletal elements as well as similar cell surface and extracellular matrix proteins^{18,19}.

Conclusions

We have identified mutations in 38 genes involved in the development and/or function of the HSN motor neurons. Mutations in most of these genes also affect other cells (see Table 1), which suggests that the HSNs are specified by a combination of genes that function during the development of multiple cell types. Furthermore, this observation indicates that there may be few genes with functions necessary only for the development of this single cell type. Similarly, there has been a recent proposal that relatively few genes function specifically in the development of the *C. elegans* touch receptor neurons.²⁰

Of the steps in the temporal pathway of HSN development (Fig. 1a), only HSN generation is not affected in any of the mutants we have studied. Perhaps we failed to identify generation mutants because we have not saturated for genes that can mutate to result in HSN functional defects³. Alternatively, genes required for the cell lineages that generate the HSNs may be essential (we would not have isolated lethal or sterile mutants in our screens) or redundant (we might not have isolated mutants that require multiple mutations for their phenotypes).

Our assignments of genes to specific HSN developmental steps depend on the assumption that the mutant phenotypes we have observed are similar to the phenotypes that result when gene activity is eliminated. As the loss-of-function phenotypes

of many of these genes have not been established, it is possible that further analyses will reveal that the assignments of certain genes must be revised.

On the basis of the mutant phenotypes and gene interactions described above, we propose a genetic pathway for HSN development (Fig. 3). One striking aspect of this pathway is that HSN development does not occur as a series of dependent steps, as might be expected from the temporal pathway (Fig. 1a). Instead, the genetic pathway reveals that much of HSN development occurs in parallel, independent steps. For example, HSN serotonin expression is not prevented by abnormalities in migration or axonal outgrowth; similarly, HSN axonal outgrowth is not prevented by abnormalities in migration or serotonin expression (see Fig. 3 legend). The independence of the expression of specific differentiated characteristics may be a general feature of neuronal development^{20,21}.

The independence of these three steps highlights the ability of the HSN to adapt to its environment. Such plasticity, common in neuronal development of vertebrates²², can be seen, for example, in the HSNs of *egl-43* mutants. From their posteriorly displaced cell bodies, these HSNs often grow axons of twice the usual length to reach their normal destination, the nerve ring (Fig. 2g). In addition, these axons sometimes form functional synapses with the egg-laying muscles, as 10% of *egl-43* animals lay eggs proficiently (Table 2). Similarly, the HSN axons in *unc-6* and *unc-40* mutants are defective in ventral growth and follow instead a novel lateral pathway to the nerve ring (Fig. 2d).

The ability of the HSNs to respond to external cues suggests that some of the genes we have studied may be expressed not within the HSNs themselves but rather within other cells that control aspects of HSN development. In this context, it is intriguing that different genes affect the outgrowth of the HSN axons in different local environments: ventral growth, defective in *unc-6* and *unc-40* mutants, is along the lateral hypodermis; anterior growth, defective in *unc-76* and *unc-34* mutants, is within the ventral cord. Some of these genes could be expressed in cells that define these different environments.

Although HSN migration, axonal outgrowth, serotonin expression, and function can be affected independently of one another, four steps—identity, sex-specific survival, maturation and specialization—are required for the expression of multiple HSN traits. All four of these steps seem to be of a regulatory nature.

The acquisition of identity is the first cell-type specific step in HSN development. The two genes that affect this step, *ham-1* and *egl-5*, do so in different ways. In *ham-1* mutants the HSNs seem to be of mixed identity: they variably express HSN traits and may also express traits normally expressed by their sisters, the PHB neurons. In *egl-5* mutants, the presumptive HSNs are defective in the expression of all HSN characteristics; these cells may be transformed to another cell type. Mutations in the *mec-3* gene in *C. elegans*²³ and the *cut*²⁴, *sevenless*²⁵ and *boss*²⁶ genes in *Drosophila melanogaster* seem similarly to prevent the acquisition of specific neuronal identities. In addition, about 40 homoeotic genes have been identified that control the acquisition of specific cell identities during *C. elegans* development²⁷. These studies indicate that the coordinate acquisition of the potential to express all aspects of a particular cell identity may be a step in the development of many cell types.

The sex-specific survival of the HSN depends upon the *egl-1* gene, which may act to interpret a signal from the sex determination pathway. *egl-1* could represent a general class of genes that interact with the sex-determination pathway to control the fates of specific sexually dimorphic cell types. *egl-1* may well be regulatory in its function, as it controls the activities of the *ced-3* and *ced-4* genes, which are responsible for programmed cell death⁶.

The step of maturation involves the transition from an early, immature HSN cell state to a later, mature state. Mutants defective in the genes *unc-86*, *egl-45* and *sem-4* all fail in late, but

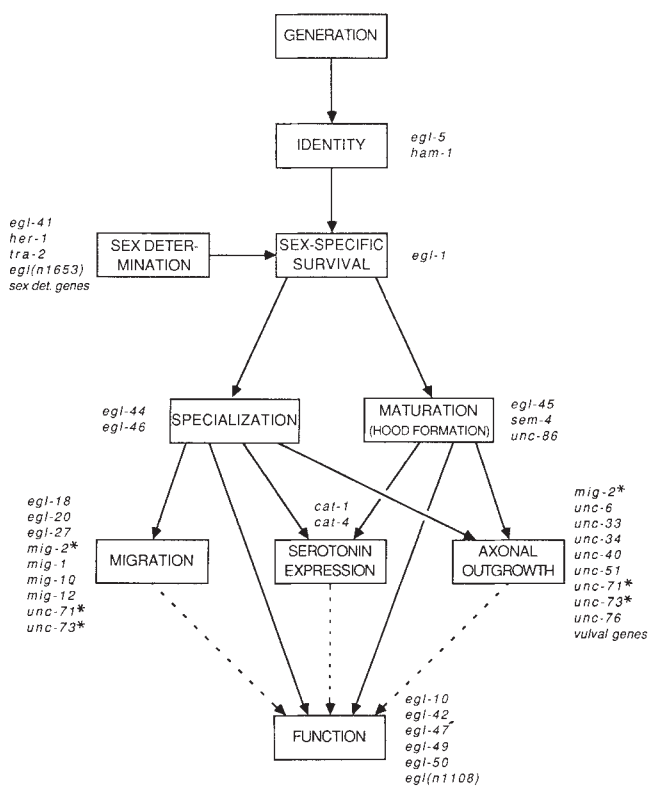


Fig. 3 A genetic pathway for the development of the HSN neurons. Boxes represent steps in HSN development. Mutations in the genes adjacent to each box affect but may not prevent that step. This pathway was constructed from the HSN phenotypes of the mutants and from a limited number of gene interaction studies. Mutants defective in a given step are defective in all steps in the pathway subsequent to that step. Hood formation has been placed with maturation because we have no mutations that distinguish this morphologically visible event from the decision step of maturation. That the migration, serotonin expression, and axonal outgrowth steps are independent is supported by the following evidence: (1) migration is normal in mutants defective in serotonin expression or axonal outgrowth; furthermore, in the wild-type, migration occurs long before axonal outgrowth and expression of detectable serotonin; (2) serotonin expression is normal in migration mutants even when the HSN cell bodies are located in the postanal region of the tail (data not shown); serotonin expression is also normal in mutants with severely attenuated axonal outgrowth, such as *unc-51* animals; (3) axonal outgrowth is completed in animals with total failures in HSN migration (data not shown) and also in *cat-1* mutants with very low serotonin levels; furthermore, axonal outgrowth occurs in the wild-type before detectable HSN serotonin appears. The dotted lines indicate that none of the mutants abnormal in HSN migration, axonal outgrowth, or serotonin expression displays the serotonin-sensitive imipramine-resistant Egl phenotype characteristic of animals with impaired HSN function. Although some of these mutants are Egl, they probably have defects in other components of the egg-laying system (data not shown). Therefore, we cannot be sure if the steps of migration, axonal outgrowth, and serotonin expression are required for HSN function. 'vulval genes' refers collectively to genes involved in vulval development (see text). 'sex det. genes', refers to other sex-determining genes, which control whether or not the presumptive HSNs undergo programmed cell death. It is likely that *mig-2*, *unc-71* and *unc-73* (denoted *), which affect both HSN migration and axonal outgrowth, function separately in these two processes (see text).

not early, aspects of HSN development. This general failure in differentiation suggests that these mutants could be blocked in a regulatory step. In support of this hypothesis, mutations in *unc-86* and *sem-4* also cause homoeotic transformations in the fates of certain cell types (ref. 28 and M. Stern, personal communication). The recent molecular cloning of *unc-86* has

revealed that this gene encodes a homoeodomain-containing protein with remarkable sequence similarity to three mammalian transcription factors²⁹. Thus, the *unc-86* protein is probably a transcription factor that controls the expression of the genes responsible for HSN differentiation, namely those genes involved in hood formation, axonal outgrowth, serotonin expression, and function. The *egl-45* and *sem-4* genes might encode other transcriptional regulators that act coordinately with the *unc-86* protein.

We propose that there is an additional regulatory step in HSN development. We call this step 'specialization', to reflect the hypothesis that it controls the expression of certain specialized HSN traits. Mutants defective in the genes *egl-44* and *egl-46* are abnormal in HSN migration, axonal outgrowth, serotonin expression, and function, but not in the acquisition of identity, sex-specific survival or hood formation. As discussed above, the fact that these genes affect multiple HSN traits without affecting similar traits in other cells suggests that these genes are regulatory. We have not included in this step three other genes—*egl-43*, *ham-2* and *ham-3*—that affect HSN migration, serotonin expression, and function, because we do not know if these genes affect axonal outgrowth (see above).

Specialization is like maturation in that it controls the expression of certain HSN traits. However, unlike the maturation genes, which control an obvious late temporal transition in the HSN cell state, the specialization genes control both early and late traits expressed at intermittent times during HSN development. The maturation and specialization genes control overlapping but distinct sets of HSN traits—only maturation controls hood formation, only specialization controls migration,

and both control axonal outgrowth, serotonin expression, and function (Fig. 3)—further exemplifying the use of parallel pathways in HSN development, even at the level of decision steps.

The genetic pathway of HSN development provides a model for the determination and differentiation of any cell type. First, a decision is made concerning whether or not a particular cell identity is to be acquired. Genes that control the acquisition of cell identity regulate other control genes, responsible for the expression of overlapping subsets of cell-specific traits. These genes, in turn, regulate the genes that encode the products responsible for specific differentiated characteristics. Further genetic and molecular analyses of the genes involved in HSN development should reveal the mechanisms responsible for each of the steps we have identified in the specification of this cell type.

We thank Nancy Tsung for assistance with the maintenance of strains; the members of our laboratory for strains, advice, and comments on the manuscript; Ed Hedgecock and Jim Manser for strains and for sharing unpublished data with us; Thomas Jessell and Jane Dodd for advice and assistance in developing immunocytochemical techniques; Nichol Thomson for help with electron microscopy; and Marty Chalfie and Jonathan Hodgkin for comments on the manuscript. This work was supported by USPHS research grants and by the Howard Hughes Medical Institute. C.D. was supported by a USPHS training grant; G.G. was supported by a postdoctoral fellowship from the Damon Runyon-Walter Winchell Cancer Fund, S.M. was supported by the Medical Scientist Training Program, Harvard Medical School.

Received 17 October; accepted 7 November 1988.

- Sulston, J. & Horvitz, R. *Dev Biol.* **82**, 41–55 (1981).
- Trent, C., Tsung, N. & Horvitz, R. *Genetics* **104**, 619–647 (1983).
- Desai, C. & Horvitz, R. *Genetics* (in the press).
- White, J., Southgate, E., Thomson, N. & Brenner, S. *Phil. Trans. R. Soc. Lond.* **314**, 1–340 (1986).
- Briley, M. in *Neuropharmacology of Serotonin* (ed R. Green) 50–70 (Oxford University Press, Oxford, 1985).
- Ellis, H. & Horvitz, R. *Cell* **44**, 817–829 (1986).
- Trent, C., Wood, W. & Horvitz, R. *Genetics* **120**, 145–157 (1988).
- Hodgkin, J. & Brenner, S. *Genetics* **86**, 275–287 (1977).
- Hodgkin, J. *Genetics* **96**, 649–664 (1980).
- Doniach, T. *Genetics* **114**, 53–76 (1986).
- Hedgecock, E., Culotti, J., Hall, D. & Stern, B. *Development* **100**, 365–382 (1987).
- Brenner, S. *Genetics* **77**, 71–94 (1974).
- Hedgecock, E., Culotti, J., Thomson, N. & Perkins, L. *Dev Biol.* **111**, 158–170 (1985).
- Trent, C. thesis, Massachusetts Institute of Technology (1982).
- Ferguson, E. & Horvitz, R. *Genetics* **110**, 17–72 (1985).
- Ferguson, E., Sternberg, P. & Horvitz, R. *Nature* **326**, 259–267 (1987).
- Desai, C. thesis, Massachusetts Institute of Technology (1988).
- Singer, S. & Kupfer, A. A. *Rev. cell Biol.* **2**, 337–365 (1986).
- Bronner-Fraser, M. *Dev Biol.* **117**, 528–536 (1986).
- Chalfie, M. & Au, M. *Science* (in the press).
- Valles, A. M. & White, K. J. *Neurosci.* **6**, 1482–1491 (1986).
- Purves, D. & Lichtman, J. *Principles of Neural Development* 117–120 (Sinauer, Sunderland, 1985).
- Way, J. & Chalfie, M. *Cell* **54**, 5–16 (1988).
- Bodmer, R. *et al. Cell* **51**, 293–307 (1987).
- Tomlinson, A. & Ready, D. *Science* **231**, 400–402 (1986).
- Reinke, R. & Zipursky, S. L. *Cell* **55**, 321–330 (1988).
- Horvitz, R. in *The Nematode Caenorhabditis Elegans* (ed W. Wood) 157–190 (Cold Spring Harbor Laboratory, 1988).
- Chalfie, M., Horvitz, R. & Sulston, J. *Cell* **24**, 59–69 (1981).
- Finney, M., Ruvkun, G. & Horvitz, R. *Cell* (in the press).
- Sulston, J., Dew, M. & Brenner, S. *J. comp. Neurol.* **163**, 215–226 (1975).
- Horvitz, R. & Sulston, J. *Genetics* **96**, 435–454 (1980).
- Hodgkin, J. *Genetics* **103**, 43–64 (1983).
- Chalfie, M. & Sulston, J. *Dev Biol.* **82**, 358–370 (1981).
- Collins, J., Saari, B. & Anderson, P. *Nature* **328**, 726–728 (1987).
- Sulston, J. & Horvitz, R. *Dev Biol.* **56**, 110–156 (1977).
- Sulston, J., Schierenberg, E., White, J. & Thomson, N. *Dev Biol.* **100**, 64–119 (1983).
- Horvitz, R., Chalfie, M., Trent, C., Sulston, J. & Evans, P. *Science* **216**, 1012–1014 (1982).

The role of the leucine zipper in the fos–jun interaction

T. Kouzarides & E. Ziff

Department of Biochemistry and Kaplan Cancer Center, New York University Medical Center, New York, New York 10016, USA

Mutagenesis of the fos protein supports the hypothesis that a heptad repeat of leucine residues stabilizes the interaction between the fos and jun proteins. We show that the complex between fos and jun can bind to DNA more tightly than either protein alone and that basic residues adjacent to the leucine repeat of fos contribute to the DNA-binding potential of the complex.

THE products of the two nuclear proto-oncogenes *c-fos* and *c-jun* form a non-covalent association *in vivo*^{1–3}. These two proteins, along with the yeast transcription factor GCN4 and the members of the myc family, share limited amino-acid homology⁴. It has been hypothesized that a heptad repeat of

leucine residues present within this region of homology may facilitate interactions between monomers of these proteins⁵. To test this hypothetical 'leucine zipper' model, we asked whether the leucine repeat region of *fos* expressed *in vitro* would be sufficient to form a complex with *jun*, and if so, whether mutation

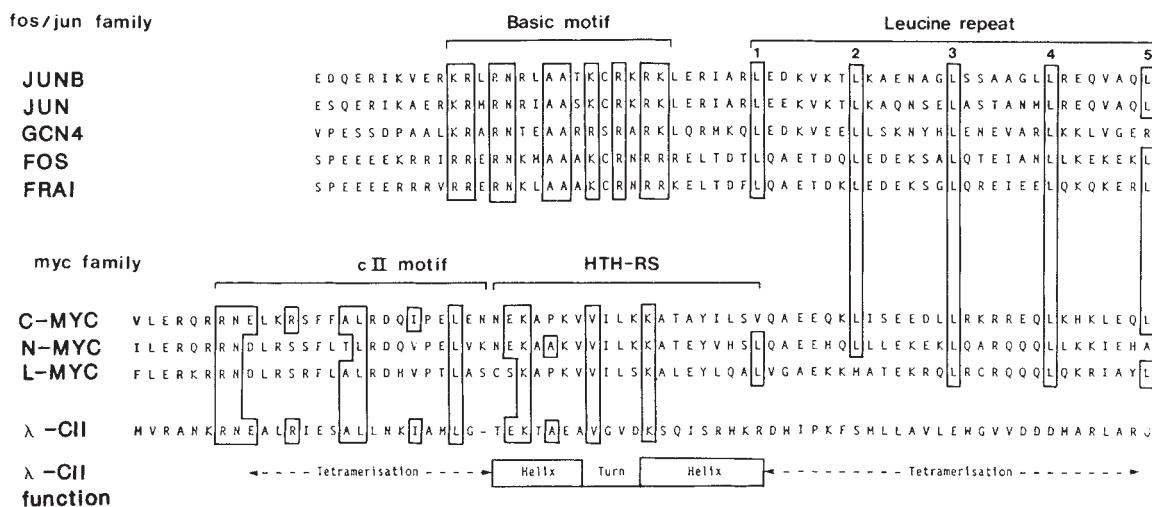


Fig. 1 a, The fos/jun and myc families of homologous proteins. Alignment of amino-acid homology between jun B¹², c-jun⁷, GCN4¹³, c-fos¹⁴, and fra 1¹⁵ starting at residue 261, 257, 133 and 103 respectively (fos/jun family). The highly conserved, mainly basic residues proposed to be part of the DNA-binding domain (basic motif) are boxed, as are the highly conserved leucine residues which form the leucine repeat. The leucine residues of the repeat are numbered 1–5. Alignment of c-myc¹⁶, N-myc¹⁷ and L-myc¹⁸ beginning from residues 361, 388 and 292 respectively (c-myc family). The conserved leucine residues that form the leucine repeat also found in the fos/jun family are boxed. Below the myc family is an alignment of the cII protein of bacteriophage λ ¹⁹ with boxes indicating exact identities. The first residue of cII is its initiator methionine. The residues of cII implicated in DNA binding (through a helix–turn–helix motif) and tetramerization are indicated below the cII sequence.

of the leucine residues would interfere with this interaction. Here we show that a substantially truncated portion of the fos protein containing the leucine repeat structure is capable of forming a complex with the jun protein. Mutagenic analysis of the leucine residues of fos supports the hypothesis that the leucine repeat structure stabilizes interactions between fos and jun. We also show that the fos–jun complex can bind to the TPA (12-*O*-tetradecanoylphorbol-13-acetate)-responsive element (TRE) more tightly than either fos or jun alone. Mutagenesis of basic residues adjacent to the leucine repeat structure of fos does not affect complex formation with jun but destroys the DNA-binding potential of the fos–jun complex. This provides evidence that the fos protein contributes to the DNA-binding potential of the fos–jun complex. We also discuss the implications of a homology between myc and the cII protein of bacteriophage λ .

Fos and jun associate *in vitro*

Alignment of the homologous domains of GCN4, jun, fos and two related proteins jun B and fra 1 (fos/jun family; Fig. 1) reveals that the most highly conserved residues within this region of homology may be arranged into two motifs. The first, the basic motif, lies at the beginning of the domain. The strongly basic nature of this motif suggests that for GCN4 (ref. 6) and c-jun⁷, this region constitutes the binding site for DNA as no other known DNA-binding structure (such as a zinc finger or a helix–turn–helix motif) is present. The second conserved motif in this domain consists of a repeat of five leucine residues, which are spaced at seven-amino-acid intervals (the leucine repeat).

Because the fos and jun proteins exist in a complex *in vivo* we sought to establish whether the heptad repeat of leucines stabilized this interaction. We first asked whether a substantially truncated portion of the fos protein containing the leucine repeat would be capable of forming a complex with the jun protein. To express such a portion of the fos protein *in vitro*, we used site-directed mutagenesis to introduce a methionine codon with a perfect Kozak consensus sequence for translation initiation just before the region of jun homology and a stop codon just following the homology (Fig. 2a). The mutagenesis also introduced two unique *Bgl*III sites whose cleavage released a restric-

tion fragment containing the truncated fos coding sequence. This fragment was cloned into an SP6 expression vector, transcribed *in vitro* using SP6 RNA polymerase and the resulting RNA was translated in a rabbit reticulocyte lysate. The expected ³⁵S-labelled 92-amino-acid translation product, called 'fos-core', is shown in Fig. 2b.

The 92-amino-acid fos-core product has three features of interest. (1) Within it are the five leucine residues proposed to be the site of interaction between fos and jun; (2) it contains the basic motif thought to be the DNA-binding site for jun and GCN4; and (3) it has the epitope for an antibody raised against a fos fusion protein antigen. We asked whether this portion of the fos protein expressed *in vitro* would form a complex with the full-length *in vitro* translation product of c-jun (Fig. 2b). Figure 3a shows that an anti-fos antibody can specifically recognize fos-core (lane 2) but not jun (lanes 3 and 11). However, it can precipitate the jun product from a mixture of fos-core and jun that has been incubated for 30 min to allow complex formation (lane 4). These data are consistent with the notion that the leucine repeat structure is the site of interaction between fos and jun. To prove the direct participation of the leucine residues in the formation of a fos–jun complex we used site-directed mutagenesis to specifically mutate individual leucines within the leucine repeat structure of fos-core (Fig. 2c). Of the five leucines in fos-core (designated leu 1 to leu 5 from the N to C terminus), leu 2 was changed to isoleucine, leu 3 was changed to arginine, and leu 4 was changed to isoleucine to generate mutants L2, L3 and L4 respectively. Two double-leucine mutants were also generated. One, L2/3, has leu 2 and leu 3 changed to isoleucine and arginine respectively and the other, L3/4, has leu 3 and leu 4 changed to arginine and isoleucine respectively. Finally, to establish whether the basic motif of fos contributes directly to the DNA-binding potential of the fos–jun complex, the two arginine residues at the C terminus of the basic motif were simultaneously changed to glutamic acid. This mutant (fos-core RR) would also serve as a control for mutants in the leucine repeat region as the altered residues fall outside, but directly adjacent to, the region of the leucine repeats and should not therefore affect the formation of the fos–jun complex.

The six mutant fos-core products were expressed *in vitro* in

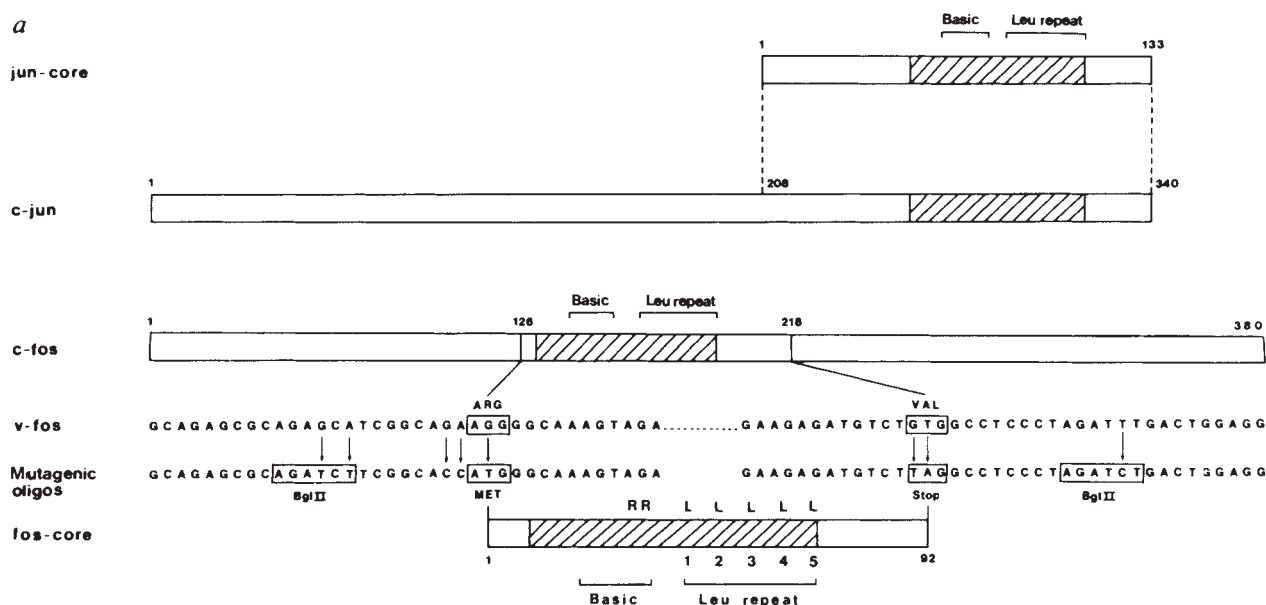
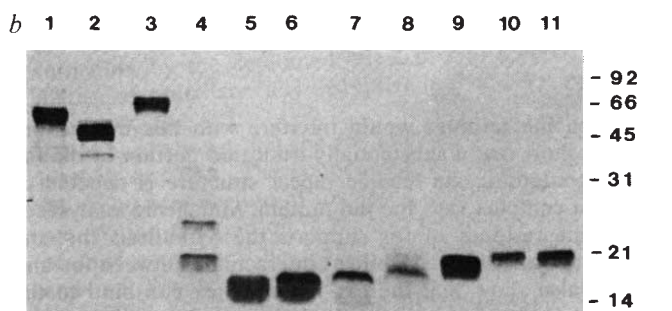


Fig. 2 *a*, Structure of fos and jun proteins. Empty boxes represent the amino-acid sequence of fos and jun and shaded boxes represent the region of homology between fos and jun shown in Fig. 1. The basic motif and leucine-repeat structure within the region of homology is shown. The complements of the mutagenic oligonucleotides used to generate fos-core coding sequence from the v-fos sequence are indicated. The changes directed to express fos-core are also shown. The two arginines at the end of the basic motif mutated in fos-core RR as well as the five leucines that make up the leucine repeat shown in Fig. 1 are indicated. The jun-core region is the DNA-binding domain of c-jun⁷. *b*, *In vitro* translation products. ³⁵S-labelled *in vitro* translation products of fos¹, jun², myc³, jun-core⁴, fos-core⁵, fos-core RR⁶, fos-core L2⁷, fos-core L3⁸, fos-core L4⁹, fos-core L2/3¹⁰ and fos-core L3/4¹¹ are shown, resolved by SDS-PAGE on a 15% polyacrylamide gel. Relative molecular masses are shown in thousands at the right. *c*, Mutations introduced within fos-core. The boxed residues indicate the substitution made in each of the six fos-core mutants. The two arginines mutated in fos-core RR are the last two residues of the fos basic motif, which are boxed in Fig. 1. At the right is a summary of two assays presented in the text: the ability of the mutant fos-core proteins to form a complex with jun (Fig. 3*a*) and the DNA-binding capability of mutant fos-core-jun complex (Fig. 5).

Methods. The mutagenic strategy was as described²⁰. To express c-fos, a fos cDNA from rat²¹ was cloned into the EcoRI site of pGEM3. The v-fos BglII fragment engineered to express fos-core was cloned in the BamHI site of pSP65 whereas the v-fos BglII fragments carrying mutations in fos-core were cloned in the BamHI site of pGEM1. Expression of jun-core was from pHJ19 (ref. 7), jun was expressed from a c-jun clone in pBSSK (a gift of R. Turner and R. Tjian) and myc was expressed from pSP65 0/1 c-myc²². *In vitro* transcription of the above constructs (1 µg each) was carried out as previously described²³. One-quarter of the RNA resulting from the transcription reaction was translated in a rabbit reticulocyte lysate in the presence of [³⁵S]methionine as described by the manufacturer (Promega).



<i>c</i>	Basic	Leu repeat					Complex with jun	DNA binding of complex
		1	2	3	4	5		
fos-core	RR	L	L	L	L	L	++	++
fos-core L2		L	I	L	L	L	++	+
fos-core L3		L	L	R	L	L	++	+
fos-core L4		L	L	L	I	L	++	++
fos-core L2/3		L	I	R	L	L	-	-
fos-core L3/4		L	L	R	I	L	-	-
fos-core RR	EE	L	L	L	L	L	++	-

the presence of [³⁵S]methionine (Fig. 2*b*) and then assayed for their ability to form a complex with the full-length c-jun product. Figure 3*a* shows that mutation of individual leucines to isoleucine (mutants L2 and L4) does not affect complex formation (lanes 5 and 7), nor does the more drastic substitution of arginine (L3; lane 6). However, the combined mutation of either of two pairs of leucine residues (L2/3 and L3/4) eliminates the capacity of fos-core to form a complex with jun (lanes 8 and 9). In contrast, mutation of residues outside the leucine repeat region (RR) does not affect complex formation (lane 10). These data provide direct evidence that the leucine residues within the

region of the leucine repeat of fos participate in interactions that stabilize the fos-jun complex.

As other nuclear proteins have been shown to possess a leucine repeat, we asked whether the interaction between the leucine repeat region of fos and jun was specific. Using the fos immunoprecipitation assay we examined whether the fos-core product was capable of forming a complex with *in vitro*-generated c-myc product (Fig. 2*b*) which also contains a leucine repeat structure (Fig. 1). The result (Fig. 3*b*) shows that fos-core is incapable of forming a complex with myc under conditions identical to those used to show interaction between fos and jun.

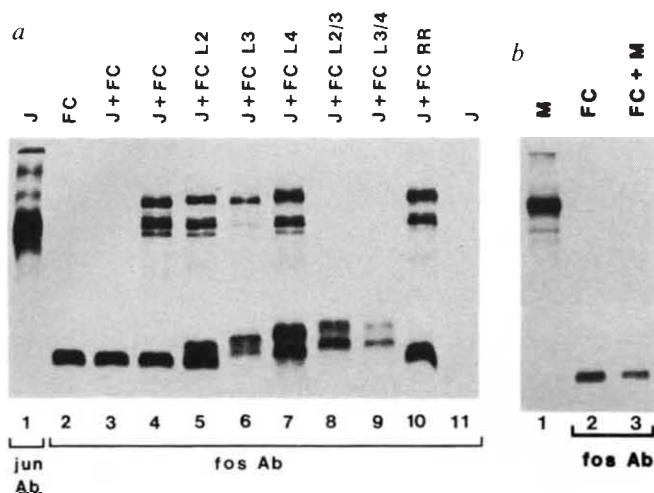


Fig. 3 The fos-core protein forms a complex with jun but not myc. *a*, The [35 S]-labelled *in vitro*-translated jun protein (J) was precipitated with the antibody (Ab) raised against a β -galactosidase fusion of the C-terminal 133 amino acids of c-jun⁷ (lane 1). The [35 S]-labelled *in vitro*-translated fos-core product (FC) was precipitated with an antibody raised against a trpE fusion to amino acids of v-fos corresponding to amino acids 73–152 of the c-fos protein²⁴ (lane 2). When the fos-core protein is mixed with the jun protein and then precipitated immediately with the fos antibody, only the fos-core protein is precipitated (lane 3). However, when fos-core and jun are mixed and then incubated at 37 °C for 30 min the fos antibody precipitates not only fos-core but also the jun protein (lane 4) indicating that fos-core and jun can form a non-covalent complex in solution. In lanes 5–10 the jun protein was mixed and incubated at 37 °C with fos-core proteins carrying mutations of leucines in the repeat structure (lanes 5–9) or mutations in the basic motif (lane 10), and then precipitated with a fos antibody. Lane 11 shows that the jun protein is not directly recognized by the fos antibody. The multiple forms of jun and fos may reflect posttranslational modifications, which can occur in the reticulocyte lysate (ref. 21 and our unpublished results). *b*, The fos-core protein was added to *in vitro*-generated myc protein (M) and the mixture incubated at 37 °C for 30 min. In lane 3, the fos antibody precipitates only the fos-core protein and not the myc protein from this incubation mixture. Lane 1 shows the 35 S-labelled translation product of myc and lane 2 shows the immunoprecipitation of the fos-core protein.

Methods. 35 S-labelled *in vitro*-translated proteins in 300 μ l RIPA buffer (10 mM Tris pH 7.4, 150 mM NaCl, 1% Nonidet P-40, 0.1% SDS, 1 mM EDTA, 10 mM KCl) were incubated at 4 °C for 15 min in the presence of 50 μ l of a 50% protein A-Sepharose slurry. After removing the Sepharose beads, 3 μ l of fos antibody was added and the mixture was incubated for 60 min at 4 °C. After the addition of 50 μ l 50% protein A-Sepharose slurry and incubation at 4 °C for 15 min, the Sepharose beads were washed four times with RIPA buffer, twice with RIPA/1M NaCl buffer and finally with RIPA buffer. The Sepharose beads were resuspended in sample buffer (15% glycerol, 2% SDS, 80 mM Tris pH 6.8, 5% β -mercaptoethanol, 0.01% bromophenol blue) and incubated in a boiling water bath for three minutes. After removing the beads, 50% of the sample was loaded on a 15% polyacrylamide gel and analysed by SDS-PAGE. The gel was then fixed in 50% methanol, 10% acetic acid, soaked in Enhance for 1 h, dried and visualized by autoradiography.

Fos and jun bind DNA cooperatively

Nuclear protein complexes containing both jun and fos have been shown to bind the consensus DNA sequence T-G-A-C-T-C-A (the TRE⁸). We therefore asked whether the fos protein expressed *in vitro* in a rabbit reticulocyte lysate translation system would bind to an oligonucleotide with the sequence of the 'fat-specific element' (FSE) which contains a TRE, and has been shown to bind the *in vivo* fos-jun complex⁸. We did not detect any binding of fos to the 32 P-labelled FSE oligonucleotide

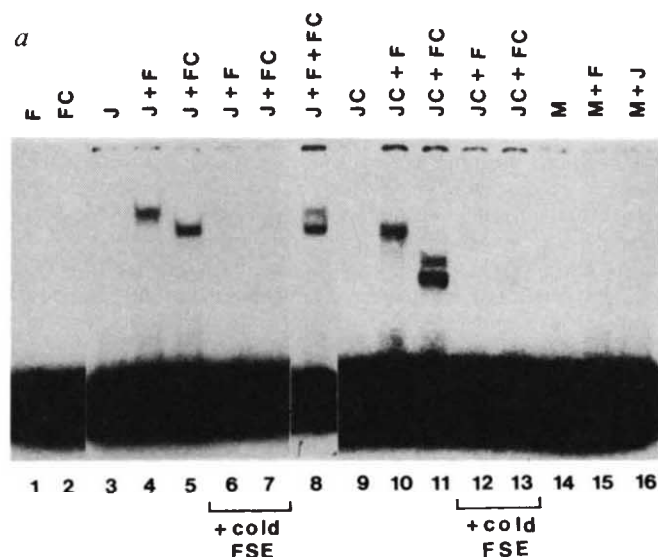


Fig. 4 *a*, Fos and jun proteins cooperate to bind DNA. Mobility shift analysis of unlabelled, *in vitro*-translated products of fos (F), fos-core (FC), jun (J), jun-core (JC), myc (M) or a combination of these products, all incubated with a 32 P-labelled oligonucleotide with the structure of the 'fat-specific element' (FSE; GATCCATGACTCAGAGGAAAACATACG) which contains a TRE element¹¹. *b*, The fos-core and jun gel shift is abolished by both fos and jun antibodies. The gel shift induced by adding fos-core and jun to the FSE (lane 1) is disrupted by either a fos antibody or a jun antibody but not by a control antibody (directed against a neuronal intermediate filament protein; D. Kegler, J. Gorham and E. Z., unpublished results). The antibodies were added during the incubation of the proteins with the FSE.

Methods. 0.2 ng 32 P-labelled FSE oligonucleotide was mixed with 2 μ l from a 50- μ l *in vitro* translation reaction, 3 μ g (3 μ l) poly(dI-dC) 800 ng (2 μ l) unlabelled, single-stranded FSE, 19 μ l binding buffer (20 mM HEPES, 10% glycerol, 100 mM KCl, 0.2 mM EDTA) and, where indicated, 300 ng unlabelled, double-stranded FSE oligonucleotide. The mixture was incubated at 37 °C for 30 min and then loaded immediately onto a 4% non-denaturing polyacrylamide gel.

(Fig. 4a, lane 1). We were also unable to detect binding of *in vitro*-generated jun to the same oligonucleotide (lane 3). However, if both the fos and jun proteins were added together in the presence of the FSE oligonucleotide, retardation of the 32 P-labelled FSE was observed (lane 4) and this retardation was specifically competed by the addition of excess unlabelled FSE (lane 6). Thus, although mixing of fos and jun proteins does not change their absolute levels, it greatly increases their binding affinity for DNA. This demonstrates that the jun-fos complex is capable of binding to the FSE more tightly than either fos or jun alone. We assume the inability to detect binding of jun alone in this assay reflects the low levels of protein generated in the rabbit reticulocyte lysate system, a limitation which does not apply to the bacterially expressed jun protein used in the DNA-binding assay of Bohmann and colleagues⁷. The fos-core protein is also capable of cooperating with jun to bind specifically to the FSE (compare lanes 2, 3, 5 and 7). The use of fos-specific and jun-specific antibodies demonstrates that retardation of the FSE oligonucleotide is dependent on the presence of both the fos-core and jun proteins, as either antibody can disrupt

retardation of the FSE, whereas a control antibody has no effect on the formation of the protein-DNA complex (Fig. 4b).

As expected, the FSE oligonucleotide is more severely retarded by the fos-jun complex (Fig. 4a, lane 4) than by the fos-core-jun complex (lane 5) as the fos protein is larger than the fos-core product. We took advantage of this difference in mobility to ask whether the fos-jun complex binds to the FSE as a dimer, tetramer or higher order multimer. If the DNA-bound complex is a heterodimer ($\alpha\beta$), composed of a single monomer of fos and a single monomer of jun then only two mobility shifts will be generated when the FSE is incubated simultaneously with jun, fos and fos-core: one shift will be due to binding of a fos-jun complex and the other due to binding of a fos-core-jun complex. If the DNA-bound complex is composed of a heterotetramer ($\alpha_2\beta_2$) however, consisting of two monomers of jun plus two monomers of fos, then an additional intermediate mobility shift should be apparent. This intermediate shift would result from the binding of a protein complex composed of two jun monomers associated with one fos monomer and one fos-core monomer. The simultaneous incubation of jun, fos and fos-core with the FSE oligonucleotide only generates two bands of shifted mobility (Fig. 4a, lane 8). No intermediate band is detected even when the two shifted bands are better resolved as a result of longer electrophoresis (data not shown). This suggests that the DNA-bound fos-jun complex is composed of a fos-jun heterodimer. However, as this assay is dependent on DNA binding, we cannot exclude the possibility that a tetramer consisting of a jun dimer complexed to fos plus fos-core is capable of forming, but is unable to bind DNA.

Having established that the 92-amino-acid fos-core region of fos was capable of cooperating with the full-length jun protein to bind DNA, we asked whether fos or fos-core would cooperate with the 133-amino-acid C-terminal portion of jun, a fragment which we term jun-core (Fig. 2a and b). The jun-core product, defined as the DNA-binding region of jun⁷ contains the leucine repeat region of jun as well as the basic motif (Fig. 1). As seen in Fig. 4a, the jun-core product (lane 9) like fos (lane 1) and fos-core (lane 2) is not capable of binding to the ³²P-labelled FSE oligonucleotide. However, when jun-core is allowed to complex with fos or fos-core, binding to the FSE is observed (lanes 10 and 11). This complex specifically competes with added unlabelled FSE oligonucleotide (lanes 12 and 13). Thus, the 92-residue fos-core and 133-residue jun-core products represent independently folding domains of fos and jun which can cooperatively bind to the FSE more tightly than either protein alone. This cooperativity can be shown to be specific to these two proteins as the myc protein, which also contains a leucine repeat and is incapable of binding to the FSE (lane 14), does not provide the cooperative DNA-binding function to either fos or jun (lanes 15 and 16).

DNA binding requires complex formation

Fos-core proteins with mutations within the leucine repeat structure (Fig. 2c) were assayed for their ability to cooperate with jun in binding to the FSE oligonucleotide. Figure 5 shows that fos-core mutants carrying mutations in two out of the five leucines (L2/3 and L3/4) are unable to bind DNA in the presence of jun. As these mutants are also incapable of forming a complex with jun (Fig. 3a), these data suggest that fos-jun complex formation is essential for DNA binding. Individual mutations of leucine residues however, reveal that the five leucines of the repeat structure are not equivalent with respect to their contribution to the DNA-binding capacity of the fos-jun complex. Mutation of leu 2 and leu 4 to isoleucine has no effect on complex formation with jun (Fig. 3a, lanes 5 and 7), yet the mutant complex L4-jun is fully competent in DNA binding (Fig. 5), whereas the mutant complex L2-jun shows a dramatic decrease in its ability to bind the FSE (Fig. 5). Mutation of leu-3 to arginine also causes a decrease in the DNA-binding capacity of fos-core-jun complex (Fig. 3a, lane 6).

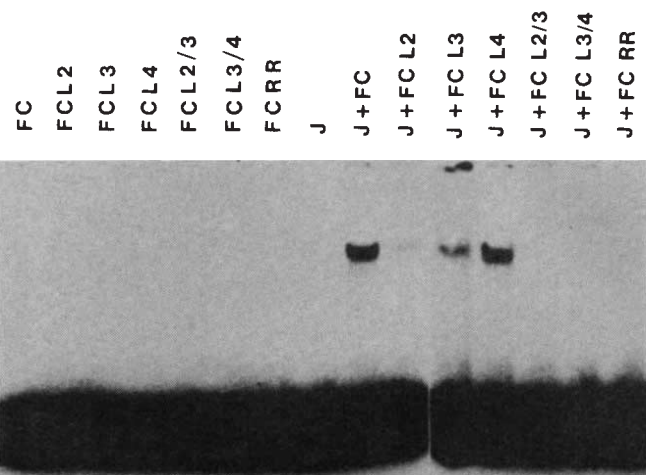


Fig. 5 DNA-binding analysis of complexes between fos-core mutants and jun. Mobility shift analysis using *in vitro* translated fos-core protein or mutants of fos-core (see Fig. 2c) incubated with ³²P-labelled FSE either alone or in combination with the jun protein (J). Methods as in Fig. 4.

Fos contributes to DNA binding

As suggested above, a basic motif directly N-terminal to the leucine repeat structure is conserved in GCN4 and jun, and is likely to be the DNA-binding motif of these proteins. We argued that conservation of this motif in the fos protein indicates that fos binds DNA through this motif when complexed to jun. We therefore asked whether a mutation of two basic residues of the fos basic motif would disrupt the DNA-binding potential of the fos-jun complex. As seen in Fig. 5, mutation of the last two arginine residues of the basic motif of fos-core to glutamic acids (fos-core RR) completely abolishes the DNA-binding capacity of the fos-core-jun complex. As this fos-core RR mutant is still fully capable of forming a complex with jun (Fig. 3a, lane 10), the fos-core product must be contributing to the DNA-binding ability of the fos-core-jun complex.

Discussion

A number of DNA-binding proteins possess a conserved heptad repeat of leucine residues. This may represent a widely used motif for the generation of a hydrophobic surface through which DNA-binding proteins interact. Mutagenic analysis of the fos leucine repeat structure presented here (summarized in Fig. 2c), provides evidence that the integrity of the leucines is essential for the interaction between fos and jun, in support of the 'leucine zipper' model⁵. The mutagenesis suggests that the relatively long hydrophobic backbone generated by the fos leucine repeat contacts the jun leucine repeat with enough strength to overcome mutation of a single leucine residue. Disruption of complex formation by the combined effect of two leucine mutations, neither of which are fully disruptive individually, supports this argument (Fig. 2c).

We have shown that the fos and jun proteins cooperate in binding to DNA. The cooperating functions reside within a 92-amino-acid portion of fos and a 133-amino-acid portion of jun. Within this portion of the two proteins is a heptad repeat of leucine residues that stabilizes the interaction between fos and jun. We believe this interaction affects the stability of a basic DNA-binding motif present directly adjacent to the leucine repeat. The fact that mutation of leucines closer to the basic motif (leu 2) has a deleterious effect on DNA binding compared with mutation of leucines further away (leu 4), supports this hypothesis. Mutation of residues within the basic motif of fos abolishes the DNA-binding potential of the fos-jun complex (Fig. 2c). This suggests that fos contributes directly to DNA

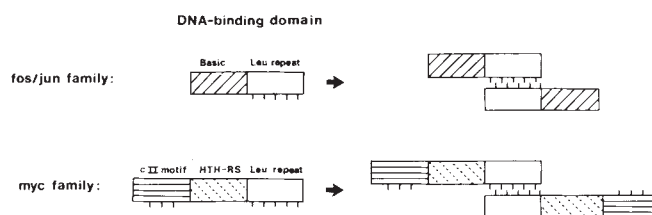


Fig. 6 The fos/jun family and the myc family may possess a DNA-binding domain with similar structural and functional organization. Data presented here establish that the DNA-binding domains of fos and jun interact through their leucine-repeat region to form a complex capable of binding to DNA through a basic motif. Homology between myc and the bacteriophage λ cII protein (Fig. 1) suggests that the myc family may possess a HTH-RS adjacent to a leucine-repeat structure, as well as an additional set of leucines (cII motif) to the N-terminal side of the HTH-RS. This arrangement of myc may be similar to that in the fos/jun family, in that protein-protein interactions through the leucine repeats may stabilize a DNA-binding domain capable of interacting with DNA.

binding and that this basic motif is the DNA-binding structure of the fos/jun family.

The presence of a DNA-binding structure next to a leucine repeat may be a common theme in DNA-binding proteins. The myc protein family, which contains a heptad leucine repeat, does not have the basic motif present in the fos/jun family adjacent to it (Fig. 1). Instead we have noticed that, at a position equivalent to the basic motif, the myc proteins show substantial homology with the DNA-binding domain of the cII protein of bacteriophage λ (Fig. 1). Extensive mutagenic analysis of the cII protein⁹ has identified a helix-turn-helix (HTH) DNA-binding structure and two domains involved in cII tetramer formation (cII function, Fig. 1). Part of the homology between

cII and the myc family of proteins falls within the HTH DNA-binding region of cII, suggesting that the myc proteins possess a HTH related structure (HTH-RS). Homology between myc and cII extends N-terminal to the HTH-RS, a region we refer to as the cII motif. Within this region, the myc proteins have three conserved leucine residues separated by an interval of seven amino acids, suggesting that, as in cII, this region of myc is involved in protein-protein interactions. In the light of the homology to cII, we suggest that the myc family of proteins possesses a potential DNA-binding domain adjacent to heptad leucine repeats, a functional organization analogous to the DNA-binding domain of the fos/jun family (Fig. 6). As the cII homologous region of the myc proteins also shows homology to the products of *MyoD1*, *AS-CT4* and *AS-CT5*^{10,11}, our speculations about myc structure may extend to encompass these proteins.

Recent reports have shown that the TRE-specific stimulation of transcription induced by the jun protein is enhanced by the presence of the fos protein^{2,3}. In the light of the data presented here, this cooperativity between fos and jun may be explained by the ability of the fos-jun complex to bind to the TRE more tightly than the jun protein alone. The fos-jun complex may have a more stable DNA-binding structure as a result of the strong interaction between the leucine repeat regions of fos and jun. As proteins related to fos and jun show extensive conservation in the DNA-binding domain (Fig. 1), interaction between other members of the fos and jun family may also result in cooperative affinity for DNA. This type of regulation at the level of protein-protein interaction may generate a diversity of protein complexes which can regulate the expression of a variety of genes.

We thank R. Turner, D. Bohmann and R. Tjian for gifts of jun plasmids and antibody, A. Klug and A. Lesk for helpful discussions. We thank the WHO and the Kaplan Cancer Center for fellowship support for T. K., and the NIH for research grants.

Received 3 October; accepted 2 November 1988.

1. Rauscher III, F. J. *et al. Science* **240**, 1010-1016 (1988).
2. Chiu, R. *et al. Cell* **54**, 541-552 (1988).
3. Sassone-Corsi, P., Lamph, W. W., Kamps, M. & Verma, I. M. *Cell* **54**, 553-560 (1988).
4. Vogt, P. K., Bos, T. J. & Doolittle, R. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3316-3319 (1987).
5. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. *Science* **240**, 1759-1764 (1988).
6. Hope, I. A. & Struhl, K. *Cell* **46**, 885-894 (1986).
7. Bohmann, D. *et al. Science* **238**, 1386-1392 (1987).
8. Rauscher III, F. J., Sambucetti, L. C., Curran, T., Distel, R. J. & Spiegelman, B. M. *Cell* **52**, 471-480 (1988).
9. Ho, Y.-S., Wulff, D. L. & Rosenberg, M. *Genes Dev.* **2**, 184-195 (1988).
10. Davis, R. L., Weintraub, H. & Lassar, A. B. *Cell* **51**, 987-1000 (1987).
11. Villares, R. & Cabrera, C. V. *Cell* **50**, 415-424 (1987).
12. Ryder, K., Lau, L. F. & Nathans, D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1487-1491 (1988).
13. Hinnebusch, A. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6442-6446 (1984).
14. Van Straaten, F., Muller, R., Curran, T., van Beveren, C. & Verma, I. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3183-3187 (1983).
15. Cohen, D. R. & Curran, T. *Molec. cell Biol.* **8**, 2063-2069 (1988).
16. Watson, D. K., Psallidopoulos, M. C., Samuel, K. P., Dalla-Favera, R. & Papas, T. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3642-3645 (1983).
17. Stanton, L. W., Schwab, M. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1772-1776 (1986).
18. Leogouy, E. *et al. EMBO J.* **6**, 3359-3366 (1987).
19. Schwarz, E., Scherer, G., Hobom, G. & Kossel, H. *Nature* **272**, 410-414 (1978).
20. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488-492 (1985).
21. Curran, T., Gordon, M. B., Rubino, K. L. & Sambucetti, L. C. *Oncogene* **2**, 80-84 (1987).
22. Hann, S. R., King, M. W., Bentley, D. L., Anderson, C. W. & Eisenmann, R. N. *Cell* **52**, 185-195 (1988).
23. Melton, D. A. *et al. Nucleic Acids Res.* **18**, 7035-7056 (1984).
24. Vosatka, R. J., Hermanowski-Vosatka, A., Metz, R. & Ziff, E. B. *J. Cell. Physiol.* (in the press).

Enhanced protein thermostability from designed mutations that interact with α -helix dipoles

H. Nicholson, W. J. Becktel & B. W. Matthews

Institute of Molecular Biology, and Departments of Physics and Chemistry, University of Oregon, Eugene, Oregon 97403, USA

Two different genetically engineered amino-acid substitutions designed to interact with α -helix dipoles in T4 lysozyme are shown to increase the thermal stability of the protein. Crystallographic analyses of the mutant lysozyme structures suggest that the stabilization is due to electrostatic interaction and does not require precise hydrogen bonding between the substituted amino acid and the end of the α -helix.

THE study of mutations that enhance the stability of proteins is of considerable practical and theoretical interest. In some instances, genetic screens have allowed the selection of mutant proteins that are more stable than their parent¹⁻⁴. In other cases

increased stability has been obtained by rational modification of the protein structure⁵⁻¹⁴. General methods of increasing stability are, however, still needed.

Recent evidence has shown that the stabilities of helical

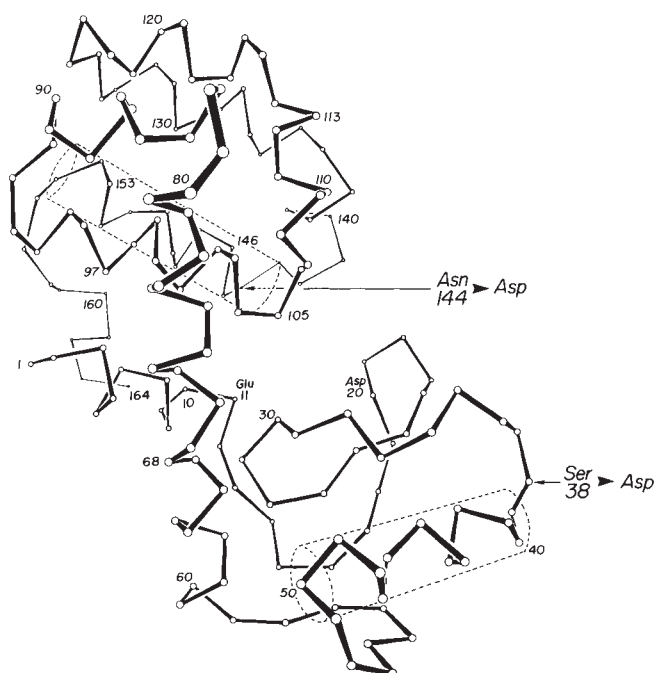


Fig. 1 Backbone structure of T4 lysozyme showing the locations of the mutations described in the text. The relevant α -helices are outlined in a broken line.

synthetic polypeptides are influenced by changes in charge at the ends of the helix¹⁵⁻¹⁷. This result has been exploited to construct a semisynthetic ribonuclease of increased thermostability⁷. It also provides a rationalization for the high frequency with which negatively-charged amino acids are observed near the N terminals of helices in known protein structures^{18,19}.

Here, we describe two successful attempts to increase the thermostability of T4 lysozyme by introducing charged residues at sites designed to interact with α -helix dipoles^{18,20,21}. In this initial study, only aspartic-acid substitutions at helical N terminals were tested. Arginine, lysine and glutamic acid residues have longer side chains than aspartic acid and their conformations are harder to predict. Also, because the $C^\alpha-C^\beta$ bonds in a helix project away from the helix axis and toward the N terminus it may be more difficult to find potentially useful substitution sites at the C terminals of α -helices. These questions will need further study.

Selection of substitutions

In the following discussion the interface residues that are half in and half out of a helix are defined as N-cap and C-cap¹⁹. It is at these residues that the backbone conformation breaks away from being α -helical. When we define the extent of an α -helix it is from residues N-cap + 1 to C-cap - 1.

There are 11 helices in T4 lysozyme²²⁻²⁴, here identified by the letters A-K. Seven of these helices were rejected from consideration because they already have negatively charged groups near their N terminals in the native lysozyme structure. For the remaining four helices, B, G, I and J, model building with the interactive computer program FRODO²⁵ was used to select aspartic-acid substitutions that would cause minimal perturbation of the three-dimensional structure of the protein.

Helix I (residues 137-141) was ruled out for several reasons. Its N terminus is very close to the C terminus of helix H (residues 126-134). Ser 136, the N-cap residue of helix I is also the C-cap residue of helix H, and makes bridging hydrogen bonds between the ends of the two helices (133 O...Ser 136 O...HN 139). These two helices are, in effect, self-stabilizing, so that the introduction of a charged residue between them would be

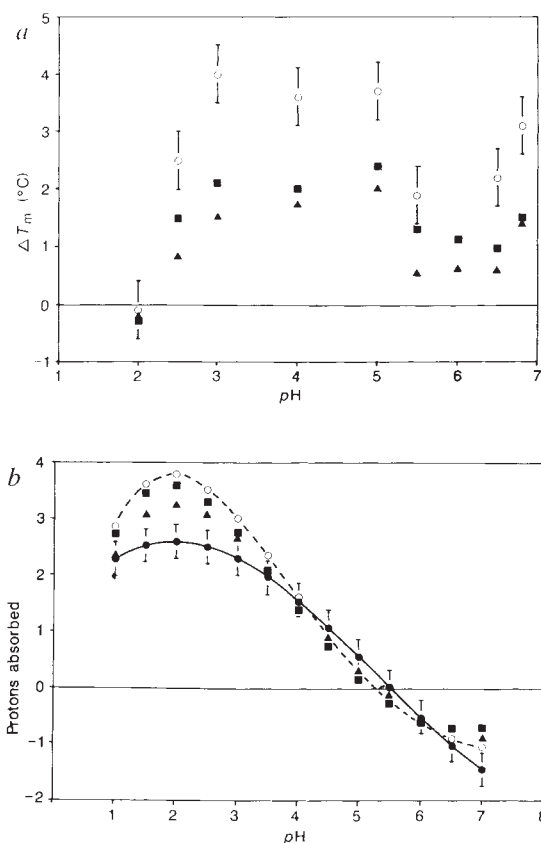


Fig. 2 *a*, Difference in melting temperature (ΔT_m) between three mutant T4 lysozymes, S38D (■), N144D (▲), double mutant (○), and the wild-type protein as a function of pH. Thermal denaturations were carried out as described³⁴. The solutions were 0.15 M in KCl and 10 mM in phosphate below pH 3 and above pH 5. Acetate replaced phosphate between pH 3 and 5. The error in melting temperature (shown only for the double mutant) was $\pm 0.5^\circ\text{C}$ or less. *b*, The difference in the uptake or release of protons upon denaturation as a function of pH for the T4 lysozyme wild-type (●; error bars included) and mutants 238D (■), N144D (▲) and S38D/N144D (○-○). The value for each protein, at each value of pH was computed as described^{33,35}. The fundamental relationship^{35,36} between the enthalpy, melting temperature and pH is $\partial T_m/\partial \text{pH} = (-2.303RT^2/\Delta H(T))\Delta\nu$ where $\Delta H(T)$ is the enthalpy of the transition and $\Delta\nu$ is the difference between the charge of the unfolded and the folded protein. The temperature dependence of the enthalpies (ΔC_p) was determined to be 2.0 ± 0.2 kcal per mol in each instance. The derivative $\partial T/\partial \text{pH}$ was obtained as the slope, at the melting temperature, of T against pH. The error in computing $\Delta\nu$ ranged from ± 0.25 to ± 0.5 protons.

counterproductive. In any case an aspartic-acid replacement at position 136 would not allow retention of these two hydrogen bonds, and would introduce unfavourable steric contacts.

It seemed that each of the three remaining helices, B (39-50), G (115-123) and J (143-155) could potentially accommodate aspartic-acid replacements at their N terminals. Helix G was not chosen initially because aspartic-acid replacements would be close to crystal contacts and might prevent crystallization. Ser 38 at the N-cap position of helix B and Asn 144 at the N-cap + 2 position in helix J were chosen as test cases for the introduction of aspartic acid. The two helices are located, respectively, in the N- and C-terminal lobes of the molecule (Fig. 1).

Mutant lysozymes Ser 38 $\rightarrow$ Asp (S38D), Asn 144 $\rightarrow$ Asp (N144D), and the double mutant S38D/N144D were obtained by oligonucleotide-directed mutagenesis^{25,26}. Procedures for DNA sequencing, cloning and protein purification were as described elsewhere²⁷⁻³¹.

Table 1 Data processing and refinement statistics

Protein	Wild-type	S38D	N144D
Cell dimensions			
<i>a</i> , <i>b</i> (Å)	61.2	61.1	61.1
<i>c</i> (Å)	96.8	97.0	97.0
Unique reflections	19,200	15,856	12,707
Resolution (Å)	1.7	1.7	1.85
<i>R</i> _{merge} (%)	9.1	4.5	5.5
<i>R</i> (%)	16.7	15.3	15.3
Δ_{Bond} (Å)	0.018	0.016	0.016
Δ_{Angle} (°)	2.2	2.2	2.4

*R*_{merge} is the agreement between intensities measured on different films. *R* is the crystallographic residual for the refined structure. Δ_{Bond} and Δ_{Angle} give the root mean square deviations of the refined bond lengths and angles from their 'ideal' values. Data for wild-type lysozyme are from ref. 24 and unpublished results. The refined coordinates for the mutant lysozymes will be deposited in the Brookhaven Data Bank.

Activity and stability

The activities of S38D, N144D and the double mutant, measured at pH 7.0 and 4 °C by the standard turbidity assay³², are respectively 320%, 160% and 430% that of wild-type lysozyme. An increase in activity has been observed for other mutations that reduce the net positive charge on the molecule (unpublished observations) and will be discussed elsewhere.

Phage T4 lysozyme can be unfolded reversibly under controlled conditions^{33,34}. Denaturations of the four proteins in this study were monitored with the change in dichroism at 223 nm, as has been described³⁴. The difference in melting temperature between the single and double mutants and the wild-type lysozyme, as a function of pH, is shown in Fig. 2a. At pH 2 there is little difference between the wild-type and mutant proteins. As the pH increases, however, both the single and the double mutant proteins melt at higher temperatures than the wild-type. The greatest differences in melting temperatures are between pH 3 and pH 5, where S38D and N144D melt approximately 2 °C higher than the wild-type. The double mutant, in which both residues 38 and 144 have been changed to aspartic acid, shows an additive effect with an approximate 4 °C increase over the wild-type melting temperature. This corresponds to an estimated³⁵ increase of 1.6 kcal mol⁻¹ in $\Delta\Delta G$, the free energy of stabilization of the double mutant relative to wild-type lysozyme.

The changes in melting temperature shown in Fig. 2a may be explained from a knowledge of the thermodynamic properties of these proteins as a function of pH and of temperature. The enthalpies of denaturation, $\Delta H(T)$, at different values of pH, were determined from the melting curves³⁴. Plots of ΔH against the melting temperature are linear, as previously observed for the wild-type and other variants of T4 lysozyme^{14,33,36}. Although the magnitude of ΔH differs for each lysozyme, its temperature dependence is the same to within the limits of experimental error³³. Given knowledge of the pH dependence of the melting temperature plus the temperature dependence of ΔH , it is then possible to compute the number of protons taken up or given off during the denaturation of each protein at different pH values^{35,36}. Although the errors associated with these calculations are large, a change of more than one proton is significant. The results for the wild-type and mutant proteins are given in Fig. 2b. Each of the Asp insertions increases the number of protons taken up on denaturation in the pH range ~1.0–4.0. This suggests that the aspartic-acid residues at sites 38 and 144 may have anomalous *pK_a*s. The magnitude of the observed stabilization corresponds to a lowering of the *pK_a* of the aspartate in the folded form of each mutant by about 0.4 relative to the *pK_a* in the unfolded form. At pH values of about 4.0 and

Table 2 Hydrogen-bonding geometry at residue 38

Hydrogen bond	Hydrogen bond length (Å)		Hydrogen bond angle (°)	
	N...O	H...O	N—H...O	H...O—C
Wild-type				
40NH...O ^γ Ser38	2.88	2.18	126.0	122.7
41NH...O ^γ Ser38	2.97	2.02	156.3	107.5
S38D				
40NH...O ^{γ1} Asp38	3.14	2.56	116.2	74.3
41NH...O ^{γ1} Asp38	2.96	2.29	123.2	97.0

The hydrogen positions were calculated from the heavy-atom coordinates as in ref. 46.

above, the aspartate side chain(s) remain charged in both the folded and unfolded forms. The folded form is stabilized by the interaction of the charged aspartate with the helix dipole, accounting for the enhanced thermostability of the mutant proteins relative to wild-type. At low pH (less than about 2.0) the aspartate side chains remain uncharged in both the folded and unfolded forms, as is the case for wild-type lysozyme, and the stabilities of the mutant and wild-type enzymes become essentially the same.

Structures of mutant lysozymes

Crystals of S38D and N144D were obtained under conditions similar to those used for the wild-type enzyme^{23,24}. Before data collection the crystals were equilibrated with a standard mother liquor containing 1.05 M K₂HPO₄, 1.26 M NaH₂PO₄, 0.23 M NaCl and 1.4 mM β-mercaptoethanol, pH = 6.7. X-ray diffraction data were collected by oscillation photography^{24,37} and the mutant structures refined with the TNT program³⁸ (Table 1).

In the map showing the difference in electron density between S38D and wild-type T4 lysozyme (Fig. 3a), there is a strong positive feature due to the extra bulk of the aspartic acid. There is also a negative feature indicating the displacement of a solvent molecule (#333) that is hydrogen-bonded to Ser 38 in the wild-type enzyme. Fig. 3b shows the overlay of the refined structures of S38D and the wild-type enzyme. Ser 38 makes hydrogen bonds with the amide nitrogens of residues 40 and 41 in the first turn of the helix. Similar hydrogen bonds are retained in the mutant structure. However, as shown in Table 2, the geometry of these hydrogen bonds in the S38D structure is much poorer than in the wild-type structure. The only other significant difference between the respective structures is a movement of the carbonyl oxygen of residue 38 coupled with the outward movement of the side chain in the mutant structure (Fig. 3b).

In the case of the mutant N144D, the difference map (Fig. 4a) shows positive and negative features indicating that the aspartic-acid replacement is slightly closer to the helix axis than the wild-type asparagine, perhaps in response to interaction with the helix dipole. The adjacent negative feature at the position labelled Cl 168 is due to the displacement of a bound chloride ion. That this is a chloride ion and not bound solvent was shown unequivocally by equilibrating lysozyme crystals in mother liquor containing NaBr in place of the standard NaCl and collecting a high resolution data set (R. Jacobson, H.N. & B.W.M., unpublished results). The largest feature (16 standard deviations) in the difference map between wild-type NaBr crystals and wild-type NaCl crystals clearly showed the replacement of a chloride ion by a bromide ion at the end of the helix. This site was previously identified as 'SOL 168' in the wild-type structure²⁴ and is therefore now identified as 'Cl 168'. The refined structure of N144D (Fig. 4b) indicates that two partially occupied water molecules (#168 and #525) take the place of the chloride ion bound to wild-type lysozyme and form hydrogen

Fig. 3 *a*, Difference in electron density between S38D mutant lysozyme and wild-type, superimposed on the structure of the wild-type enzyme. Coefficients are $(F_{\text{Mut}} - F_{\text{WT}})$ where F_{Mut} and F_{WT} are the observed amplitudes, and the phases are from the refined model of wild-type lysozyme. Resolution is 1.7 Å. Contours drawn at $\pm 4.0\sigma$ where σ is the r.m.s. density throughout the unit cell. Positive contours solid, negative broken. *b*, Superposition of the refined structure of S38D (open bonds) on wild-type lysozyme (solid bonds). Hydrogen bonds, with distances shown in Å, are indicated by thin lines.

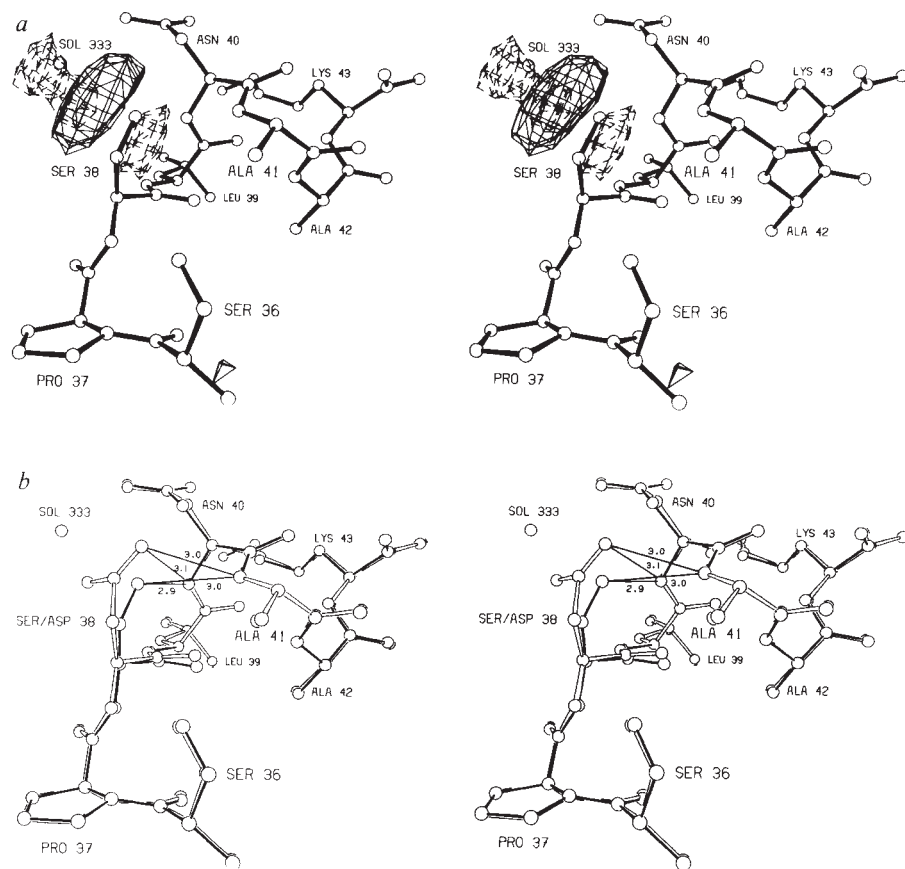
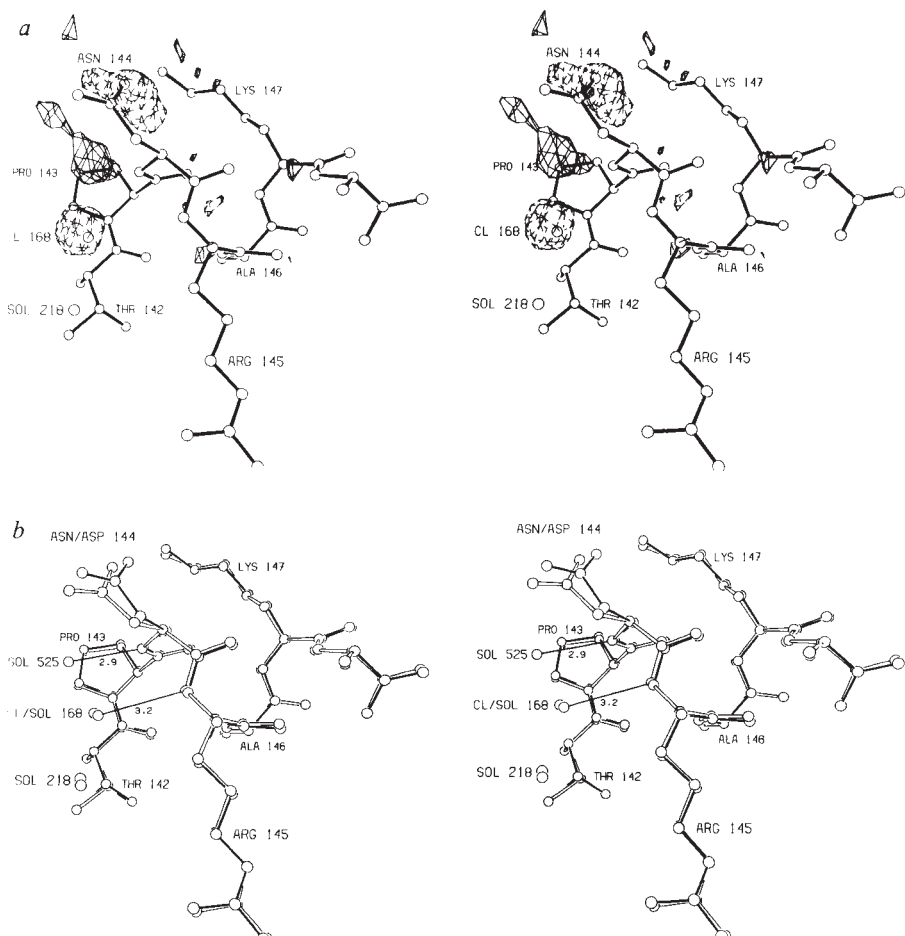


Fig. 4 *a*, Difference in electron density between N144D mutant lysozyme and wild-type. All conventions and contour levels as in Fig. 3*a* except that the resolution is 1.85 Å. Cl 168 is a chloride ion that is displaced in the mutant structure (see text). *b*, Superposition of the structure of N144D lysozyme (open bonds) on that of wild-type (solid bonds). The chloride ion, Cl 168, present in the wild-type structure, is replaced in the mutant lysozyme by solvent molecules #168 and #525 which seem to be distributed with partial occupancies between the two sites shown.



bonds to the amide nitrogens of residues 144 and 145. It should be noted that Lys 124 of an adjacent molecule in the crystal structure is within 4.6 Å of the bound chloride ion and may contribute to its binding in the crystal lattice. The role of the chloride ion in solution is therefore uncertain (there is another feature in the chloride/bromide map—six standard deviations—located at the beginning of helix G, the third helix that survived the selection procedure. No chloride ion was observed at the beginning of helix B, presumably because Ser 38 caps this helix and prevents the approach of a bound counterion). Other than the differences discussed above, the structures of N144D and wild-type lysozyme are essentially identical.

Structural basis of enhanced stability

The two mutant lysozymes described here were designed to increase thermostability by enhanced helix-dipole interactions. Both designs seem to have been successful, although the structural details are shown to be quite different.

In the first case, an unchanged helix-capping residue, Ser 38, is replaced with aspartic acid. The Asp 38 side chain makes hydrogen bonds similar to those made by Ser 38, but the geometry is poorer (Table 2). This shows that the added stability of the mutant lysozyme is not simply due to improved hydrogen bonding.

In the second substitution, Asn 144 → Asp, neither the wild-type Asn nor the replacement Asp have direct hydrogen bonds to the end of the helix. Rather, the respective side chains are located somewhat away from the helix axis (Fig. 4b). Because Asp 144 is located at the N-cap + 2 position, that is within the body of the α -helix, its side chain cannot rotate into a position to make hydrogen bonds with any of the exposed backbone amides at the end of the helix. As can be seen in Fig. 4b, rotation of the Asp (or Asn) 144 side chain about its C $^{\alpha}$ -C $^{\beta}$ bond would cause prohibitively close contacts with the amides of residues 144 and 145 and the amide of residue 143 is not available because this amino acid is a proline. The enhanced thermostability in this case clearly cannot be attributed to stronger hydrogen bonding by the Asp 144 side chain to the end of the helix. The simplest interpretation, for each replacement, is that the mutant structure is stabilized by an electrostatic interaction between the negatively-charged side chain and the positive charge attributed to the α -helix dipole^{18,20,21}. The importance of electrostatics is also strongly supported by the pH dependence of the stabilities of the mutant lysozymes relative to wild-type (Fig. 2a). As discussed above, it is only at pH values where the aspartates are negatively charged in the folded state that the mutant structures are more stable than wild-type.

One possibility that needs to be considered is stabilization by adventitious salt-bridges. This seems virtually impossible in the case of S38D as the closest positive charge is 9 Å away. In the case of N144D, the closest positive charge (N $^{\epsilon}$ of Lys 147) is 4.7 Å from the closest oxygen of Asp 144 but there are no significant features around Lys 147 in the difference map indicating movement towards Asp 144. In addition, the mutant Asp 144 side chain is actually further from Lys 147 than is the wild-type asparagine.

Another possible factor that needs to be considered is stabilization from reduction in the overall charge on the molecule³⁶. This is currently being tested in a systematic way but the similar stabilities of pairs of mutant lysozymes such as P86R and P86D (ref. 39), and T157N and T157D (ref. 40) suggest that changes in thermostability due to a single charge change in a solvent-exposed residue are relatively small.

Inferences for protein stabilization

The lysozyme mutants S38D and N144D provide two successful examples of protein stabilization by rational amino-acid replacement. α -helices are commonly found in proteins and therefore provide many potential target sites. Even though T4 lysozyme is a small protein with only 164 amino acids, it has 11 α -helices.

We show here that at least two of these can be stabilized by relatively straightforward substitutions.

The observation that protein stabilization can come from the generalized electrostatic interaction of an amino-acid side chain with an α -helix dipole should greatly simplify the design of stabilizing replacements. It is clearly easier to devise amino-acid substitutions that introduce charges near the ends of α -helices if it is not also required that such replacements hydrogen bond to the ends of the helices. There are many ways in which one can envisage introducing an aspartic or glutamic acid close to the N terminus of a helix. As one specific example, many α -helices in known protein structures have asparagine or serine residues at the N-cap position¹⁹. These provide potential sites for mutation to aspartic acid, provided there is not already a negatively-charged residue in the vicinity. Interaction of basic side chains with the C terminals of helices could provide additional routes to stabilization. This is being tested. High-resolution crystallographic analysis suggests that the loss of helix-dipole interactions at the C terminus of α -helix 82-90, together with the introduction of strain, destabilizes the temperature-sensitive mutant lysozyme Arg 96 → His⁴¹.

The results presented here, and those of others^{8,42,43}, suggests that stabilization resulting from multiple mutations is additive. Thus, the combination of several mutations, each of which may cause only a small increase in stability, can provide substantial improvement in protein stability⁴⁴.

The presence of chloride ions at the terminals of two of the helices in wild-type T4 lysozyme demonstrates that there is positive potential that is not satisfied intramolecularly at these sites. Many interactions have been tabulated between helices and counterions in several proteins^{21,45} and used as evidence for the existence of the helix dipole. The identification of helices that have bound counterions at their ends may provide a selection technique for introducing thermostable mutations. If the presence of bound ions can be determined crystallographically, as was done here, then it may be possible to design a side-chain substitution to displace the counterion. The charged side chain may replace the electrostatic role of the counterion but avoid the unfavourable entropy that is expended in binding the ion. Indeed, any sites where ions are observed to bind could be potential sites for stabilization by the introduction of appropriately charged replacement residues.

We especially thank T. Alber for establishing many of the genetic techniques used here, J. Wozniak for technical assistance, D. C. Muchmore for the expression system, R. Jacobson, R. Faber and K. Wilson for access to unpublished data, and R. L. Baldwin, J. Bell, F. W. Dahlquist, S. Dao-pin, M. Matsumura, L. McIntosh, D. E. Tronrud and L. H. Weaver for advice and helpful discussions. The work was supported in part by grants from the NIH (to B.W.M. and J. A. Schellman), NSF (to B.W.M.) and the Lucille P. Markey Charitable Trust.

Received 19 September; accepted 27 October 1988.

1. Alber, T. & Wozniak, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 747-750 (1985).
2. Shortle, D. & Lin, B. *Genetics* **110**, 539-555 (1985).
3. Matsumura, M. & Aiba, S. *J. biol. Chem.* **260**, 15298-15303 (1985).
4. Bryan, P. N. *et al. Proteins Struct., Func. Genet.* **326-334** (1986).
5. Perry, L. J. & Wetzel, R. *Science* **226**, 555-557 (1984).
6. Wetzel, R., Perry, L. J., Baase, W. A. & Becktel, W. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 401-405 (1988).
7. Mitchinson, C. & Baldwin, R. L. *Proteins: Struct. Funct. Genet.* **1**, 23-33 (1986).
8. Hecht, M. H., Sturtevant, J. M. & Sauer, R. T. *Proteins: Struct. Funct. Genet.* **1**, 43-46 (1986).
9. Sauer, R. T. *et al. Biochemistry* **25**, 5992-5998 (1986).
10. Imanaka, T., Shibasaki, M. & Takagi, M. *Nature* **324**, 695-697 (1986).
11. Villafranca, J. E., Howell, E. E., Oatley, S. J., Warren, M. S. & Kraut, J. *Trans. R. Soc.* **A317**, 405-413 (1986).
12. Wells, J. A. & Powers, D. B. *J. biol. Chem.* **261**, 6564-6570 (1986).
13. Pantoliano, M. W. *et al. Biochemistry* **26**, 2077-2082 (1987).
14. Matthews, B. W., Nicholson, H. & Becktel, W. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6663-6667 (1987).
15. Ihara, S., Ooi, T. & Takahashi, S. *Biopolymers* **21**, 131-145 (1982).
16. Shoemaker, K. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2349-2353 (1985).
17. Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M. & Baldwin, R. L. *Nature* **326**, 563-567 (1987).

18. Blagdon, D. E. & Goodman, M. *Biopolymers* **14**, 241-245 (1975).
19. Richardson, J. S. & Richardson, D. C. *Science* **240**, 1648-1652 (1988).
20. Wada, A. *Adv. Biophys.* **9**, 1-63 (1976).
21. Hol, W. G. J., van Duijnen, P. T. & Berendes, H. J. C. *Nature* **273**, 443-446 (1978).
22. Matthews, B. W. & Remington, S. J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4178-4182 (1974).
23. Remington, S. J. *et al. J. molec. Biol.* **118**, 81-98 (1978).
24. Weaver, L. H. & Matthews, B. W. *J. molec. Biol.* **193**, 189-199 (1987).
25. Jones, T. A. in *Crystallographic Computing* (ed. Sayre, D.) 303-317 (Oxford University Press, Oxford, 1982).
26. Zoller, M. J. & Smith, M. *DNA* **3**, 479-488 (1984).
27. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. *Meth. Enzym.* **154**, 367-382 (1987).
28. Alber, T. *et al. in Protein Structure, Folding and Design, UCLA Symp. Cell. Molec. Biol.* 307-318 (Liss, New York, 1986).
29. Alber, T. & Matthews, B. W. *Meth. Enzym.* **154**, 511-533 (1987).
30. Grütter, M. G. *et al. J. Molec. Biol.* **197**, 315-329 (1987).
31. Owen, J. E., Schultz, D. W., Taylor, A. & Smith, G. R. *J. molec. Biol.* **165**, 229-248 (1983).
32. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E. & Dahlquist, F. W. *Meth. Enzym.* (in the press).
33. Tsugita, A. (1971) in *The Enzymes* 3rd edn Vol. 5 (ed. Boyer, P. D.) 343-411 (Academic, New York, 1971).
34. Elwell, M. & Schellman, J. A. *Biochim. biophys. Acta* **494**, 367-383 (1977).
35. Becktel, W. J. & Baase, W. A. *Biopolymers* **26**, 619-623 (1987).
36. Becktel, W. J. & Schellman, J. A. *Biopolymers* **26**, 1859-1877 (1987).
37. Hawkes, R., Grütter, M. G. & Schellman, J. *J. molec. Biol.* **175**, 195-212 (1984).
38. Schmid, M. F. *et al. Acta Crystallogr.* **A37**, 701-710 (1981).
39. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta Crystallogr.* **A43**, 489-501 (1987).
40. Alber, T. *et al. Science* **239**, 631-635 (1988).
41. Alber, T. *et al. Nature* **330**, 41-46 (1987).
42. Weaver, L. H. *et al. Biochemistry* (submitted).
43. Matsumura, M., Yasumura, S. & Aiba, S. *Nature* **323**, 356-358 (1986).
44. Baase, W. A., Muchmore, D. C. & Becktel, W. J. *Biophys. J.* **49**, 109a (1986).
45. Matthews, B. W. *Biochemistry* **26**, 6885-6888 (1987).
46. Hol, W. G. J. *prog. Biophys. molec. Biol.* **45**, 149-195 (1985).
47. Baker, E. N. & Hubbard, R. E. *Prog. Biophys. molec. Biol.* **44**, 97-179 (1984).

LETTERS TO NATURE

A low-temperature companion to a white dwarf star

E. E. Becklin* & B. Zuckerman†

* Institute for Astronomy, University of Hawaii,
2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Department of Astronomy, University of California, Los Angeles,
California 90024, USA

We have discovered an infrared object located about 120 AU from the white dwarf GD165. With the exception of the possible brown dwarf companion to Giclas 29-38 which we reported last year¹, the companion to GD165 is the coolest (2,100 K) dwarf star ever reported and, according to some theoretical models, it should be a sub-stellar brown dwarf with a mass between 0.06 and 0.08 solar masses. These results, together with newly discovered low-mass stellar companions to white dwarfs, change the investigation of very low-mass stars from the study of a few chance objects to that of a statistical distribution. In particular, it appears that very low-mass stars and perhaps even brown dwarfs could be quite common in our Galaxy.

We are currently conducting an infrared photometric survey of about 200 white dwarf stars to search for orbiting low-mass companion stars and brown dwarfs. A brown dwarf is a sub-stellar object of mass $<0.08 M_{\odot}$, which is unable to sustain nuclear fusion reactions in its interior. Earlier studies in this survey failed to detect any cool, low-mass companions to the young white dwarfs in the Hyades and Pleiades clusters², but later we observed excess infrared (IR) radiation from the white dwarf G29-38 (ref. 1). We argued that the most natural source

of this IR emission is a 1,200-K brown dwarf in orbit around G29-38. Because its source is not yet spatially resolved, it is also possible that the excess IR emission is not produced by a brown dwarf but rather by a bizarre cloud of dust particles in orbit around, and heated to 1,200 K by, G29-38.

The present study was initiated to shed light on the question³ of whether stars and brown dwarfs with masses less than $\sim 0.2 M_{\odot}$ are rather common. We will report elsewhere (B.Z. and E.E.B., manuscript in preparation) that, in addition to G29-38 and GD165, we have discovered seven very-low-mass ($\sim 0.1 M_{\odot}$) stellar companions to white dwarfs. This represents a significant rate of detection of such objects and suggests that many objects of mass $\leq 0.1 M_{\odot}$ are floating freely between the stars. If a substantial quantity of dark unseen matter does exist in the solar vicinity^{4,5}, then our data suggest that brown dwarfs and very low-luminosity stars need to be considered seriously as its carrier. Here we describe the rather unusual properties of the companion to GD165.

Table 1 contains data on GD165 and other cool stars, obtained in late July 1988 with the 3-m NASA Infrared Telescope Facility (IRTF) at the Mauna Kea Observatory in Hawaii. Also shown are data for some of the coolest known stars obtained by other observers using other telescopes. We employed the standard IRTF InSb photometric system with apertures of various angular diameters in the telescope focal plane, although most of the data in Table 1 were obtained with an aperture diameter of 3.5 arcsec (full width at half maximum power). The secondary mirror was chopped by 15 or 20 arcsec at 7 Hz between an object of interest and a reference position, while the telescope was nodded every 10 sec so that the object appeared first in one beam and then in the other. The angular separation between GD165 and its companion GD165B (Table 1) was measured with the single detector system on the IRTF. This separation

Fig. 1 Infrared images of GD165 and GD165B obtained by K. Hodapp, J. Rayner, D. Hall and one of the authors (E.E.B.) on the UH 2.2-m telescope in August 1988. Infrared wavelengths are: a, 1.25 μm (J); b, 1.65 μm (H); c, 2.2 μm (K). The images were obtained using a new, very sensitive, Hg: Cd: Te array detector developed by Honeywell Electro-Optics Division in collaboration with the Jet Propulsion laboratory and the University of Hawaii for a second-generation infrared instrument on the Hubble Space Telescope. Each image is the sum of four 24-s exposures of the 8×8 array co-added on a half-pixel grid. The scale was 0.9 arcsec per pixel and the seeing was ~ 1 arcsec. These conditions cause the images to appear to be asymmetric. The white dwarf GD165 is the northern object. North is up and east is to the left; each frame is ~ 7 arcsec on a side.

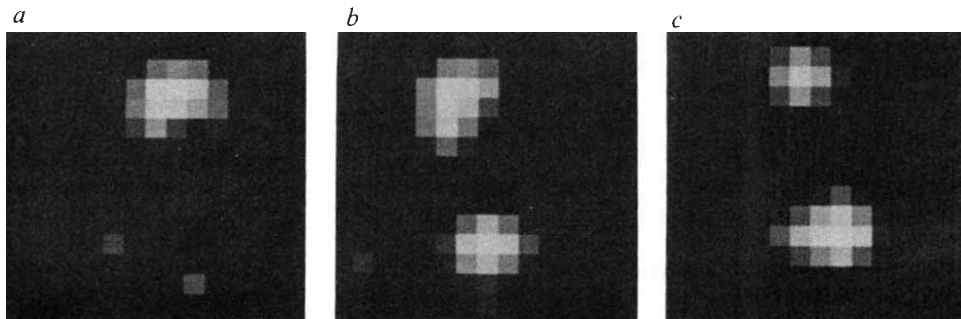


Table 1 Measured infrared brightness for low-luminosity stars

Star	α (1950)* (Epoch 1988.7)	δ (1950)* (Epoch 1988.7)	V (0.55 μm)	J (1.25 μm)	H (1.65 μm)	K (2.2 μm)	L (3.5 μm)	L' (3.8 μm)	Ref. †
GD165	14 ^h 22 ^m 12 ^s 0	9°30'47"	14.32	14.48 ± 0.04		14.48 ± 0.05			
GD165B	0°7' West	4°2' South		15.75 ± 0.10	14.76 ± 0.04	14.09 ± 0.04	13.33 ± 0.12		
LHS2924	14 ^h 26 ^m 36 ^s 1	33°24'07".5	19.74	11.92	11.18	10.73	10.25 ± 0.03		
				11.90	11.17	10.68	10.06		‡
G ℓ 569			10.20	6.60	6.02	5.78	5.58		
G ℓ 569B	1°2' East	5°0' North		10.78 ± 0.20	10.15 ± 0.15	9.65 ± 0.15	9.14 ± 0.10		
					10.16	9.56		8.93	15
VB8			16.78	9.88	9.23	8.86		8.35	16
(G ℓ 644C)				9.75	9.19	8.84	8.5		9
				9.77	9.18	8.80			17
VB10			17.8	9.90	9.24	8.80	8.5		9
(G ℓ 752B)				9.92	9.24	8.81			17

* The 1950 positions and offsets were measured at the IRTF and should be accurate to a few arcsec (positions) or one arcsec (offsets). These tabulated positions are for epoch 1988.7.

† If no reference is given, the tabulated J , H , K , L magnitudes were measured at the IRTF, where it is assumed that Vega = 0 mag at all wavelengths.

‡ Ref. 8 and R. G. Probst, personal communication.

agrees, within the errors, with that subsequently determined with an infrared array (see Fig. 1 and below).

Errors listed in Table 1 are one standard error of the mean and are entirely statistical for GD165 and GD165B and for the L magnitude measured for LHS2924. For G ℓ 569B, the errors are larger and include our best estimate of the uncertainty associated with correcting for scattered light from G ℓ 569. The tabulated errors do not include uncertainties in the photometric standards used. Some indication of uncertainties in the standards and possible systematic errors may be gained by examination of the magnitudes measured for LHS2924, VB8 and VB10 by different observers, as for these bright isolated objects statistical errors should be completely negligible. Derived properties of some of the low-luminosity stars in Table 1 are listed in Table 2.

Figure 1 shows infrared images of GD165 and GD165B obtained at 1.25, 1.65 and 2.2 μm on the University of Hawaii 2.2-m telescope on Mauna Kea by Hodapp *et al.* (see figure legend). These images verify the colours and point-like nature of the components.

There is only a small chance that GD165B is a background star or galaxy that is unrelated to the white dwarf GD165. We estimate that the probability is only ~ 0.001 that a random background star or galaxy with a K magnitude of 14 would lie within 4.3 arcsec of GD165 (refs 6, 7). Because we have examined ~ 100 white dwarfs, it might appear at first glance that the probability that GD165B is merely a very red background object is as large as 0.1. The colours of GD165B are sufficiently unusual, however, that if this were the case, then

this object would be most probably either a highly redshifted galaxy or a very distant dust-enshrouded red giant star. The expected surface density of such stars or galaxies at apparent K magnitude 14 is orders of magnitude lower than the surface density that we used to arrive at the probability of 0.001 above. In addition, the images of GD165B presented in Fig. 1 appear to be stellar rather than diffuse (which would be expected for a galaxy).

Figure 2 is a near-infrared colour-colour diagram in which we have plotted most of the lowest-luminosity stars known. According to Probst and coworkers^{8,9}, cool main-sequence stars lie in the region around the curved solid and dashed lines in the lower left-hand corner of the plot.

It is evident from Table 2 and Fig. 2 that, possibly excepting the source of the excess IR radiation at G29-38, GD165B is the coolest low-luminosity, stellar-like object ever identified outside our Solar System. According to some theoretical models for the evolution of low-mass objects¹⁰, GD165B is too cool to be a star regardless of its age; rather, it is a sub-stellar brown dwarf. If so, we estimate that its mass lies between 0.06 and 0.08 $M_{\odot}$. The lower mass limit is based on the cooling ages of white dwarfs^{11,12} and brown dwarfs¹⁰. Specifically, the time for GD165 to cool to its present temperature of $\sim 12,000$ K (ref. 13) is $\sim 6 \times 10^8$ years. All brown dwarfs with masses less than 0.06 $M_{\odot}$ would have cooled below 2,100 K after this time interval¹⁰. Given the uncertainty of $\sim 25\%$ in the distance to GD165, the radius that we derive for GD165B (Table 2) is in reasonable agreement with the brown-dwarf model¹⁰ radius of $\sim 0.08 R_{\odot}$. Alternatively, some theoretical models¹⁴ would classify GD165B as a star at

Table 2 Derived quantities for low-luminosity stars

Star	Distance D (pc)†	Projected linear separation a (AU)‡	Absolute K magnitude M_K (mag)§	Surface temperature T_* (K)	Stellar radius R_* ($R_{\odot}$)	Stellar luminosity L_* ($L_{\odot}$)	Ref. for D
GD165B¶	29	123	11.78	2,130	0.061	7.8×10^{-5}	See text
LHS2924	11.1		10.49	2,800	0.074	3.4×10^{-4}	18
G ℓ 569B	10.4	53	9.5	2,775	0.11	7.4×10^{-4}	19
VB8	6.2		9.87	3,100	0.086	7.0×10^{-4}	19
VB10	5.8		9.99	3,000	0.084	5.8×10^{-4}	19

† Excepting GD165B, distances are determined from trigonometric parallax, which should be accurate to 10%.

‡ Projected linear separation of secondaries from primaries are calculated from the distance D and the angular offsets given in Table 1.

§ K magnitude of the stars if they were located 10 pc from the Earth.

|| Estimated by fitting infrared brightnesses from 1.25 to 3.5 μm to a black-body curve.

¶ For GD165B, derived quantities are based on the very likely, but still not proven, assumption that it is a true companion of GD165. The distance to GD165 was derived from the absolute and apparent visual magnitudes given by McCook and Sion¹³, and should be accurate to $\sim 25\%$ (ref. 20).

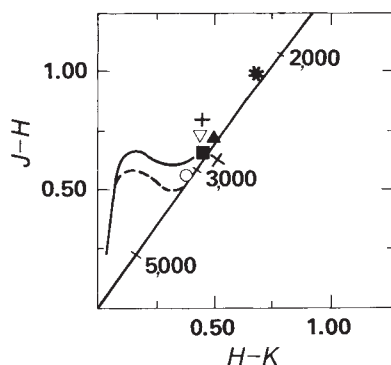


Fig. 2 An infrared colour-colour diagram for cool, low-luminosity objects. The black-body line is labelled in degrees K. According to refs 8 and 9, cool main-sequence stars lie in the region that encompasses the curved solid and dashed lines in the lower left corner. The symbols represent: *GD165B; ▲ LHS 2924; ■ VB10; ○ VB8; + GJ623B; × GJ569B; ∇ RG0050-2722. All data are from Table 1 except for GJ623B from ref. 21 and RG0050-2722 from ref. 22.

the very bottom of the main sequence (mass $\sim 0.08 M_{\odot}$), rather than as a brown dwarf.

It is amusing to speculate on how the night sky might have appeared had we lived on a planet orbiting the main-sequence predecessor of the white dwarf GD165 at the same distance that the Earth orbits the Sun. At that distant time in the past, GD165B would have been much closer to GD165, because the orbit of GD165B expanded when GD165 lost most of its mass during the giant phase of stellar evolution (see, for example, ref. 2). We expect that GD165B orbited roughly as far from the main-sequence predecessor of GD165 as Neptune and Pluto are from the Sun. If so, then GD165B, from our hypothetical Earth-like perspective, would have appeared in the night sky as a red object with about the same brightness as Mars.

In conclusion, we have discovered a cool, spatially resolved object that could be a brown dwarf in orbit around the white dwarf GD165. This result, in combination with those we have reported elsewhere, suggests that very-low-mass stars and possibly even brown dwarfs may be important constituents of our Galaxy.

We are extremely indebted to Drs Klaus Hodapp, Don Hall and John Rayner for obtaining and reducing the infrared images presented in Fig. 1. We gratefully acknowledge Drs S. Kleinmann, M. Tamura and I. Gatley for obtaining images of GD165 for us at the NOAO 50-inch telescope at Kitt Peak. We thank Drs J. Gradie, J. Greenstein and R. Probst for assistance and discussions, and Mr R. L. O'Daniel for expeditiously editing and typing the manuscript. This research was supported in part by grants from NASA and the NSF to the University of Hawaii and from the NSF to UCLA. E.E.B. and B.Z. are visiting astronomers at the Infrared Telescope Facility operated by the University of Hawaii under contract to NASA.

Received 12 September; accepted 21 October 1988.

1. Zuckerman, B. & Becklin, E. E. *Nature* **330**, 138-140 (1987).
2. Zuckerman, B. & Becklin, E. E. *Astrophys. J.* **319**, L99-L102 (1987).
3. Liebert, J. & Probst, R. G. *Astr. Astrophys.* **25**, 473-519 (1987).
4. Bahcall, J. N. *Astrophys. J.* **287**, 926-944 (1984).
5. Bienaymé, O., Robin, A. C. & Crézé, M. *Astr. Astrophys.* **180**, 94-110 (1987).
6. Chokshi, A. thesis, Univ. of California (1986).
7. Lilly, S. J. & Cowie, L. L. in *Infrared Astronomy With Arrays* (eds Wynn-Williams, C. G. & Becklin, E. E.) 473-482 (Univ. of Hawaii, Honolulu, 1987).
8. Probst, R. G. & Liebert, J. *Astrophys. J.* **274**, 245-251 (1983).
9. McCarthy, D. W., Probst, R. G. & Low, F. J. *Astrophys. J.* **290**, L9-L13 (1985).
10. Nelson, L. A., Rappaport, S. A. & Joss, P. C. *Astrophys. J.* **311**, 226-240 (1986).
11. Iben, I. & Tutukov, A. V. *Astrophys. J.* **282**, 615-630 (1984).
12. Mazzitelli, I. & D'Antona, F. *Astrophys. J.* **308**, 706-720 (1986).
13. McCook, G. P. & Sion, E. M. *Astrophys. J. Suppl.* **65**, 603-671 (1987).
14. D'Antona, F. & Mazzitelli, I. *Astrophys. J.* **296**, 502-513 (1985).
15. Forrest, W. J., Strutskie, M. F. & Shure, A. *Astrophys. J.* **330**, L119-L123 (1988).
16. Berriman, G. & Reid, N. *Mon. Not. R. astr. Soc.* **227**, 315-329 (1987).

17. Stauffer, J. R. & Hartmann, L. W. *Astrophys. J. Suppl.* **61**, 531-568 (1986).
18. McCarthy, D. W. *et al. Astrophys. J.* (in the press).
19. Gliese, W. *Catalogue of Nearby Stars* no. 22 (Veröf. des Astr. Rechen-Inst. Heidelberg, 1969).
20. Greenstein, J. L., *Publ. astr. Soc. Pacif.* **100**, 82-89 (1988).
21. McCarthy, D. W. & Henry, T. *Astrophys. J.* **319**, L93-L98 (1987).
22. Reid, I. N. & Gilmore, G. *Mon. Not. R. astr. Soc.* **196**, 15P-18P (1981).

A new interpretation of emission-like features in γ -ray burst spectra

Fulvio Melia

Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208, USA

The spectra of γ -ray bursts, which are generally thought to result from a transient phenomenon on a neutron star, extend from the optical up to γ -ray energies. It has recently been suggested¹ that the whole spectrum may be due to irradiation of a reprocessing and reflecting boundary near a source of power-law γ -rays with a low-energy cutoff $E_0 \approx 300$ –500 keV. In this picture, the emission at intermediate energy ($10 \text{ keV} \leq E \leq 300$ –500 keV) results from the reflection of incipient high-energy γ -rays by the surface layers of the underlying neutron star and/or its accretion disk. Some bursts display high-energy (~ 400 -keV) features, which may be gravitationally redshifted photons produced from the annihilation of electron-positron pairs. This explanation, however, has been open to criticism. I show that an alternative explanation for the high-energy features may be the superposition of the incipient and reflected spectral components near E_0 .

Emission-like features in γ -ray burst (GRB) spectra were first reported by Mazets *et al.*², who suggested that these lines were due to gravitationally redshifted ($z \approx 0.2$ –0.3) photons from the annihilation of electron-positron (e^-e^+) pairs. Although this initial interpretation is still open to debate, there seems to be little doubt that the features do exist. Lines in the vicinity of 400 keV have now been reported in $\sim 10\%$ of all GRBs, and have been observed simultaneously by more than one experiment in at least three cases³ (GB781104, GB781119 and GB811231).

The interpretation of high-energy features as e^-e^+ annihilation lines has proved to be problematic for several reasons. First, the derived line strength depends on the continuum subtraction method used⁴, such that the exact peak energy and linewidth are uncertain. In addition, it is difficult to reconcile the apparently large widths ($\Delta E \geq 200$ keV) of these features with the 'standard' model in which the lines are due to pair annihilations on the surface of the underlying neutron star⁵. Second, it is unclear why annihilation features should be seen in only a subset of all GRBs. Third, linewidths of absorption features (also seen in some GRB spectra) and emission features are both uncorrelated and incompatible with continuum temperatures. There is also no correlation between the emission and absorption lines themselves.

Thus, although a natural interpretation of these features is that they are redshifted e^-e^+ annihilation lines, alternative explanations have been sought. For example, Norris⁶ has suggested that the continuum spectrum at the beginning of some bursts is strongly suppressed below ~ 400 keV. If the cutoff energy drops very rapidly, the integrated spectrum would then have a line-like feature. The 'emission' lines may be dependent on the geometry, or they may simply be an artefact of the spectral deconvolution process. These features could perhaps even be due to rapidly varying two-component spectra. Kluźniak⁷ has shown that high-energy features may be due to in-flight annihilation of relativistic positrons traversing a moving plasma. If this is the case, their peak energies may be taken as evidence for a rapidly rotating underlying neutron star.

The latter scenario, however, seems to be incompatible with

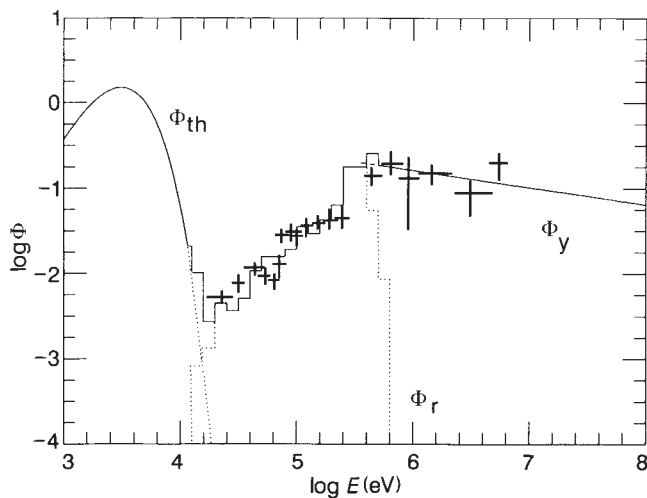


Fig. 1 The composite spectrum resulting from the irradiation of a neutron star by power-law γ -rays ($\Phi_\gamma = |\alpha| [E/E_0]^\alpha$), characterized by a spectral index $\alpha = -0.2$, low-energy cutoff $E_0 = 350$ keV, and anisotropy index $\eta = 40$ (see text). The thin dotted lines show the thermal component Φ_{th} resulting from radiative diffusion of the deposited energy, and the component Φ_r resulting from the reflection of γ -rays by the stellar surface. The incipient and composite spectra are indicated by the dashed and solid lines, respectively. Also shown for comparison are the data points corresponding to the spectrum of GB780325 (adapted from Hueter¹⁸).

the detection of a 2.2-s period in the 1984 August 5 event⁷, as well as the apparent discovery of periodicities ranging from 1 to 10 s in other GRBs⁸. Nonetheless, these period measurements, together with the confirmation of low-energy features (interpreted as cyclotron resonances) in some GRB spectra⁹, have strengthened the case for a neutron-star origin of GRBs. Evidence for this identification has also been provided by studies of the long-wavelength emission associated with some bursts¹⁰⁻¹⁶, which support the conjecture that at least part of the GRB spectrum ($E \lesssim 10$ keV) may result from reprocessing of γ -rays in a neutron-star environment. In addition, it has recently been shown that the steep spectral component at intermediate energy ($10 \text{ keV} \lesssim E \lesssim 400 \text{ keV}$) may be due to the reflection of high-energy γ -rays by the stellar surface and perhaps by a cold accretion disk¹. In this picture, the low-energy absorption features are associated with the reflected component and are due to the enhanced scattering cross-section at the cyclotron resonance that is expected in the surface layers of a magnetized neutron star. Here I explore the possibility that the superposition of the reflected component and the incipient power-law γ -rays could also account for the spectral feature seen at ~ 300 –500 keV in some bursts¹.

An important clue to the nature of GRB spectra is that both absorption and emission features, as well as the continuum spectrum, evolve significantly during the course of a burst². Emission-like features tend to be observed in strong bursts with hard spectra⁶, and usually near the burst onset¹⁷, precisely where the 'secondary' components (that is, the thermal emission from the heated stellar surface and the reflected radiation) are weakest. A classic example of a burst that displays both a low-energy absorption line and a high-energy feature is GB780325 (ref. 18) (Fig. 1). Its spectrum above 300 keV is clearly in excess of the continuum extrapolated from the 60–300-keV range. Because the detector has good resolution, is well shielded to suppress Compton-scattered events, and the model used in assembling the data does not contain a line, the observed discontinuity cannot be attributed to compliance.

The Gamma-Ray Spectrometer on the Solar Maximum Mission satellite has discovered that >1 -MeV emission (with spectral indices $-1 \lesssim \alpha \lesssim 1$) is a common component of GRBs,

and that in many cases it actually dominates the emitted power¹⁹. It is therefore not unreasonable to assume that the incipient γ -ray spectrum is dominated by power-law emission, $\Phi(E) = |\alpha| (E/E_0)^\alpha$, above some cutoff energy E_0 (perhaps near the discontinuous jump in the spectrum of GB780325). Here the emitted flux $\Phi \propto E^2 (dN/dE)$, where dN/dE is the photon number flux per energy interval. This proposition gains support from the discovery that spikes associated with the 2.2-s pulsation in the spectrum of GB840805 seem to be due to the higher-energy photons of this event⁷. Power-law spectra may be produced by various processes, such as synchrotron self-Compton, which is operating in this energy range in active galactic nuclei²⁰, or the first-order Fermi acceleration of photons by scattering in an accretion shock or in a cold converging fluid flow²¹.

To calculate the reflected component and the energy deposition profile within the heated layers of the star (and/or disk), I have carried out Monte Carlo simulations of the γ -ray transport below the source of incipient γ -rays, taking into account the angular and energy dependence of the Klein-Nishina cross-section (see ref. 1 for a more complete discussion). The γ -ray flux produced by the mechanisms described above is unlikely to be isotropic. I have therefore allowed for anisotropic emission by defining the 'anisotropy index' $\eta \equiv \langle F_{\gamma,ns} \rangle / \langle F_{\gamma,0} \rangle$, where $\langle F_{\gamma,ns} \rangle$ and $\langle F_{\gamma,0} \rangle$ are the angle-averaged flux directed toward the neutron star and the flux directed along the line-of-sight, respectively¹. Figure 1 shows the resultant, composite spectrum seen at an inclination angle $i = 80^\circ$ with respect to the symmetry axis in the case when only the stellar surface contributes to the thermal (Φ_{th}) and reflected (Φ_r) components. (This situation may arise, for example, when no disk is present, or when the Alfvén radius is significantly greater than the stellar radius R_* .) In this example, a neutron star ($R_* = 10^6$ cm) is irradiated by γ -rays with index $\alpha = -0.2$ and low-energy cutoff $E_0 = 350$ keV, originating from a point source at a distance $H_\gamma = R_*$ above the polar cap. The values of the parameters α and E_0 were chosen to roughly fit the power-law component in the spectrum of GB780325. Of the remaining parameters, only η has a significant effect on the individual components Φ_r and Φ_{th} , but then the effect is only on the magnitude of these fluxes relative to Φ_γ , and not on their spectral shape. For GB780325, the anisotropy index seems to lie in the range $20 \lesssim \eta \lesssim 40$. Values of η larger than this produce a 'hump' near 250–500 keV which is bigger than that observed, whereas smaller values of η result in a reflected component Φ_r that cannot account for the emission at intermediate energies.

Although no attempt was made to incorporate the enhanced scattering opacity at the cyclotron resonance, this simple model exhibits a remarkably good fit to the discontinuous jump near 300–500 keV and the continuum below the hump. It therefore offers a viable alternative explanation for the high-energy features seen in some bursts. I have assumed that the neutron star alone is responsible for reprocessing and reflecting some of the incipient γ -rays. Of course, if an accretion disk is present¹⁵, it too would contribute to the thermal and reflected fluxes. Indeed, the disk spectral components may be dominant in many GRB spectra¹⁶. This is not the case for GB780325, however, because the strong feature at ~ 60 keV (if interpreted as a cyclotron line) requires a significant contribution to Φ_r from the stellar surface.

Is it possible then to understand the rapid variability of these features within the context of this model? If the incipient γ -rays are indeed produced by a non-thermal process above the polar cap of the neutron star, it is conceivable that the physical parameters, such as the source height H_γ , are variable on a dynamic timescale ($\lesssim 10^{-4}$ s). This might be the case, for example, if the power-law γ -rays result from the first-order Fermi acceleration of photons by scattering in a 'standing' shock inside an accretion funnel above the stellar surface. One could speculate that the accretion rate increases during the course of the GRB, pushing the shock closer to the surface layers of the

neutron star and consequently changing the value of η . Thus, the rapid variability of the high-energy features and their general absence in the latter part of bursts may be due to a variable anisotropy index. Such a scenario is consistent with the fact that cyclotron features, which require a large contribution to Φ_r from the stellar surface, tend to appear preferentially during the main part of the burst. It may also be possible for η to be much greater than unity during portions of some bursts. The fact that not all GRBs have significant emission at high energy²² supports the expectation that the reflected γ -rays sometimes dominate the incipient γ -ray emission along the line of sight.

The next generation of instruments will allow us to observe GRBs at energies $E \leq 1$ keV. It is difficult to see how a significant reflected component Φ_r could be produced without simultaneously producing significant thermal emission at $E \sim 1$ keV, unless adiabatic cooling by expansion of the heated layers shifts the radiation out of the X-ray band. The existence or absence of a thermal component Φ_{th} could thus be an important diagnostic for testing this model.

This work was supported in part by the NSF.

Received 3 August; accepted 17 October 1988.

1. Melia, F. *Astrophys. J.* (in the press).
2. Mazets, E. P. *et al. Pisma Astr. Zh.* **6**, 706-111 (1980).
3. Liang, E. P. & Petrosian, V. *Am. Inst. Phys. Conf. Proc.* (American Institute of Physics, New York, 1986).
4. Fenimore, E. E., Klebesadel, R. W., Laros, J. G. & Stockdale, R. E. *Nature* **297**, 665-667 (1982).
5. Kluźniak, W. *Astrophys. J.* (submitted).
6. Norris, J. P. thesis. Univ. of Maryland (1983).
7. Kouveliotou, C. *Astrophys. J.* **330**, L101-L105 (1988).
8. Barat, C. *et al. Astrophys. J.* **286**, L5-L9 (1984).
9. Murakami, T. *et al. Nature* **335**, 234-235 (1988).
10. London, R. A. & Cominsky, L. R. *Astrophys. J.* **275**, L59-L63 (1983).
11. Melia, F. *et al. Astrophys. J.* **305**, L51-L55 (1986).
12. Hartmann, D. H. *et al. Bull. Am. astr. Soc.* **18**, 928 (1986).
13. Melia, F. *Fundam. cosmic. Phys.* **12**, 241-254 (1987).
14. Melia, F. *Proc. Sofia COSPAR Symp.* (ed. White, N.) (in the press).
15. Melia, F. *Astrophys. J.* **324**, L21-L25 (1988).
16. Melia, F. *Astrophys. J.* (in the press).
17. Cline, T. L. in *High Energy Transients in Astrophysics* (ed. Woosley, S. E.) 333-342 (Am. Inst. Phys., New York, 1984).
18. Hueter, G. J. in *High Energy Transients in Astrophysics* (ed. Woosley, S. E.) 373-377 (Am. Inst. Phys., New York, 1984).
19. Matz, S. M. *et al. Astrophys. J.* **288**, L37-L40 (1985).
20. Melia, F. & Königl, A. *Astrophys. J.* (in the press).
21. Kafka, P. & Meyer, F. in *High Energy Transients in Astrophysics* (ed. Woosley, S. E.) 578-580 (Am. Inst. Phys., New York, 1984).
22. Matz, S. M. thesis, Univ. of New Hampshire (1986).

Bulk synthesis of the 81-K superconductor $\text{YBa}_2\text{Cu}_4\text{O}_8$ at high oxygen pressure

J. Karpinski, E. Kaldis, E. Jilek, S. Rusiecki & B. Bucher

Laboratorium für Festkörperphysik, ETH, 8093 Zürich, Switzerland

The superconducting phase $\text{YBa}_2\text{Cu}_4\text{O}_8$ ('124') differs from $\text{YBa}_2\text{Cu}_3\text{O}_7$ ('123') in having a double, instead of single, Cu-O chain running parallel to the b axis. It was first observed as a lattice defect in partly decomposed 123 powders¹, and then as an ordered defect structure in inhomogeneous 123 thin films^{2,3}, but has not up to now been synthesized in bulk. As part of a systematic study of the pressure-temperature-composition phase diagram of pseudo-binary Y-Ba-Cu-O systems at high oxygen pressure^{4,5}, we have determined the approximate field of thermodynamic stability of the 124 phase. Here we report synthesis of the phase in bulk at 400 bar O_2 and 1,040 °C. In agreement with previous measurements in thin films, the transition temperature is 81 K. Unlike the 123 compound, the oxygen content in 124 is thermally stable up to 850 °C, which may prove important for applications.

Figure 1 shows the latest data for the pressure-temperature-composition (P - T - x) phase diagram, presented as a van't Hoff

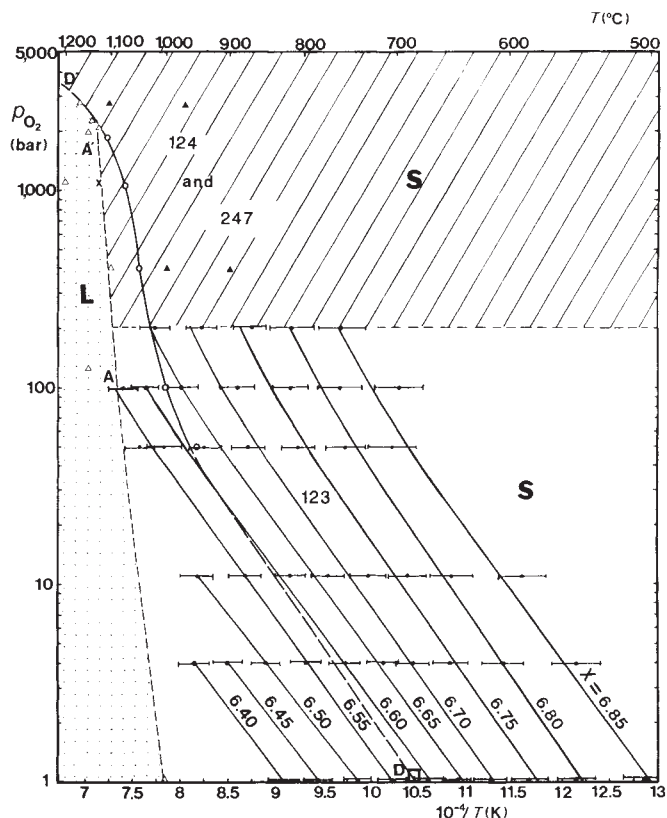


Fig. 1 Revised P - T - x phase diagram of the Y-Ba-Cu-O pseudobinary system for various oxygen compositions x . At $P \leq 100$ bar O_2 the 123 compound is stable; at higher pressures the new phases 124 and 247 appear. Curve D-D' denotes the onset of strong decomposition; A-A' is the melting curve; $\square$ denotes the orthorhombic-tetragonal phase transition for $P = 1$ bar. L and S denote regions of liquid and solid phases respectively.

diagram ($\log P$ versus $1/T$), for various values of the oxygen content x . The line A-A' shows the melt boundary and the line D-D' denotes the onset of pronounced decomposition at various oxygen pressures and temperatures. Such decomposition can be clearly resolved in our measurements (open circles) at $P \geq 50$ bar. We have argued previously⁴ that the strong decomposition coincides with an orthorhombic-to-tetragonal phase transition. For $P = 1$ bar, this transition takes place at $T = 700$ °C, and the line D-D' passes through this temperature. The linearity of both the isocompositional curves of $\text{YBa}_2\text{Cu}_3\text{O}_{7+x}$ (for $x = 6.50$ - 6.85) and the D-D' curve at $P \leq 100$ bar indicates that, up to this pressure, a single-phase region exists. X-ray and energy-dispersive X-ray analysis (EDAX) characterization show that this is the stability range of the 123 phase. At higher pressures, the slope of the D-D' curve changes in the region 100-200 bar and 2,000-3,000 bar, indicating the appearance of new phases. These phases have been detected previously^{5,6}; EDAX characterization (Fig. 1a and c of ref. 5) indicated a Cu content of ~ 3.5 and ~ 4 Cu atoms per formula unit, respectively. X-ray structural analysis on a single crystal of the former phase confirmed the formula $\text{YBa}_2\text{Cu}_{3.5}\text{O}_{7+x}$ (ref. 7). X-ray powder (Guinier) diagrams of the latter phase, corresponding to the formula $\text{YBa}_2\text{Cu}_4\text{O}_8$, could be indexed assuming a space group $Ammm$ and the atomic positions determined from X-ray analysis of a thin film⁸. Contrary to ref. 8, however, our work based on 64 reflections yielded orthorhombic lattice constants $a = 3.871$ (1) Å, $b = 3.840$ (1) Å and $c = 27.25$ (4) Å, with low standard deviations. Neutron diffraction⁹ confirmed the purity of this phase and gave refined structural parameters. Figure 2 shows the observed neutron powder diffraction pattern for $2\theta = 5$ - 70° , compared with that calculated assuming space group $Ammm$.

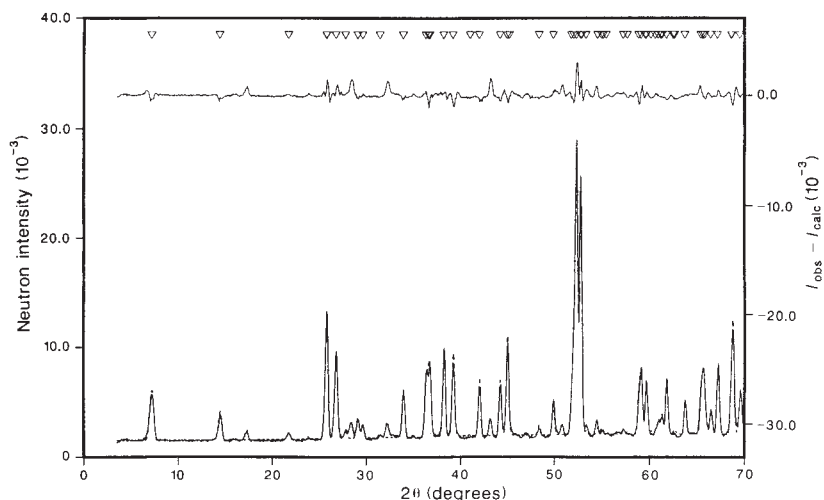


Fig. 2 Neutron diffraction spectra of $\text{YBa}_2\text{Cu}_3\text{O}_{8.043 \pm 0.001}$, for $2\theta = 2-70^\circ$ (ref. 9). With the exception of a few weak impurity lines, all reflections belong to the 124 phase. The lower curves refer to the left-hand scale; the solid line is the experimental result and the dashed line is the result of calculations, assuming space group $Ammm$ and lattice parameters as in ref. 8. The upper curve is the difference between these two curves.

and the atomic positions of ref. 8. Some weak lines from an unidentified impurity are also seen in this diffraction pattern.

Particular attention was given to the determination of the oxygen non-stoichiometry. Recently, we have developed a technique for high-accuracy volumetric analysis of the oxygen content of oxides (K. Conder, S. Rusiecki and E. Kaldis, to be published); for $\text{YBa}_2\text{Cu}_4\text{O}_{8+y}$, for example, y can be determined with an accuracy of ± 0.001 . This method is based on previous work on the determination of the non-stoichiometry of hydrides¹⁰ and involves the very accurate collection and

measurement of the gas evolved during dissolution of the oxide in an $\text{HNO}_3:9\text{H}_2\text{O}$ solution. Samples of the 124 phase showed a non-stoichiometric range $-0.150 < y < 0.100$. Of particular interest is the thermal decomposition curve of this metastable phase at 1 bar oxygen pressure. Thermobalance measurements show that $\text{YBa}_2\text{Cu}_4\text{O}_8$ decomposes rapidly only above 850°C , and after slow cooling the decomposition products are $\text{YBa}_2\text{Cu}_3\text{O}_{6.97} + \text{CuO}$ (Fig. 3a). Similar results are obtained in argon, but in that case the decomposition starts at $\sim 700^\circ\text{C}$. To investigate the kinetics of decomposition, a pure 124 sample was heated on the thermobalance at 770°C and 1 bar O_2 for 50 h. The weight of the sample decreased by 0.16% in the first 8 h but was virtually constant subsequently. Starting from a composition $\text{YBa}_2\text{Cu}_4\text{O}_{7.997}$, the final product was therefore $\text{YBa}_2\text{Cu}_4\text{O}_{7.920}$ (Fig. 3b). No additional phases were evident in the X-ray analysis. We infer, therefore, that the oxygen content of this compound is thermally stable up to $T \leq 850^\circ\text{C}$. Figure 3a and b shows clearly that the oxygen losses are not reversible under the above conditions, constituting evidence that this compound is thermodynamically metastable at normal (1 bar) pressure. Metastability results from a high activation energy of decomposition, which is due, in turn, to the increase in Cu-coordination of the oxygen in the double, edge-sharing square-planar CuO chains. This coordination change has been described for SrCuO_3 and other oxides¹¹.

The fact that metastable thin layers of this compound may be synthesized remains unexplained. It is possible that the

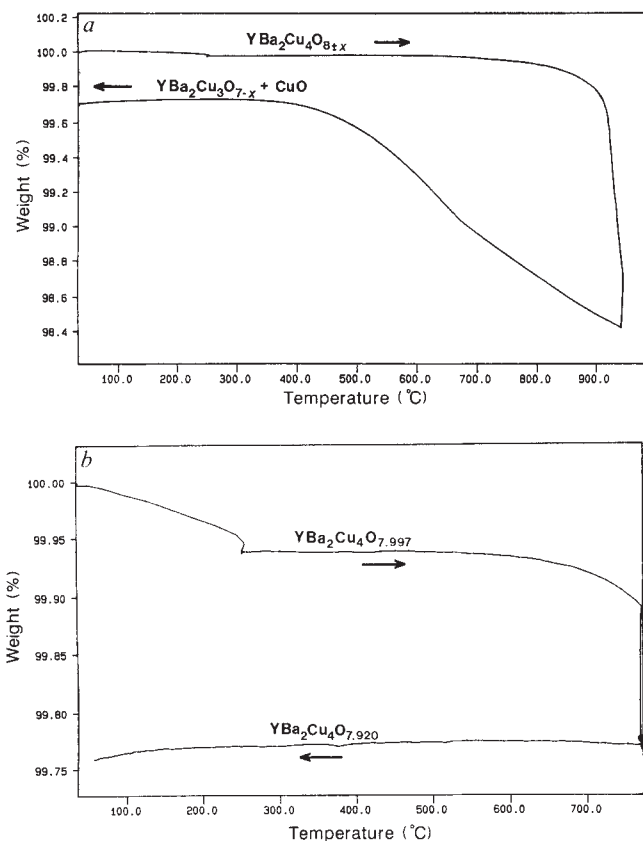


Fig. 3 *a*, Thermal (complete) decomposition curve of $\text{YBa}_2\text{Cu}_4\text{O}_8$, from thermobalance measurements. Appreciable decomposition occurs at $T \geq 850^\circ\text{C}$. Decomposition at $T > 900^\circ\text{C}$ leads to $\text{YBa}_2\text{Cu}_3\text{O}_{7-x} + \text{CuO}$. *b*, Thermal (partial) decomposition curve of a sample heated to 770°C for 50 h. During cooling no recovery of oxygen takes place. The apparent slight decrease in weight at low temperatures is due to the form of the base line.

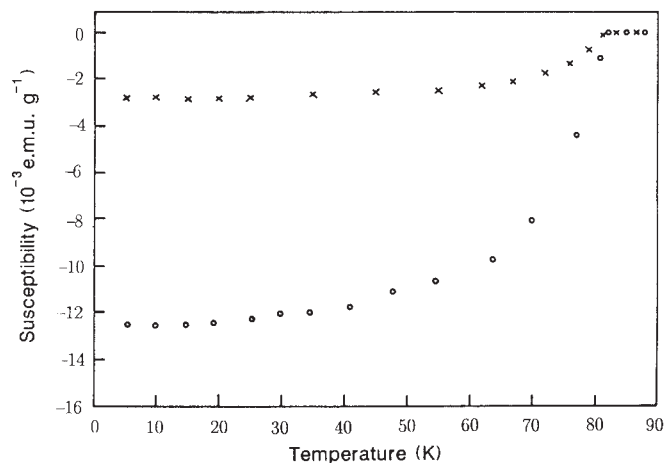


Fig. 4 Magnetic susceptibility of bulk $\text{YBa}_2\text{Cu}_4\text{O}_8$ as a function of temperature in a field of 100 G. Crosses indicate field-cooling and circles indicate zero-field-cooling. The transition temperature is $T_c \approx 81\text{ K}$.

substrate interfacial forces stabilize the thin films. However, the fact that these films are a mixture of 124 and 123 phases^{2,8,12,13} indicates that the defects at the interface of the domains of these two structures may also be responsible for the stabilization.

The d.c. magnetic susceptibility was measured with a sample-moving magnetometer in a magnetic field of 100 G (Fig. 4). A 200-mg sample was analysed under both field-cooled and zero-field-cooled conditions. The zero-field-cooled run revealed a nearly complete Meissner effect. A Meissner effect of ~20% was deduced from field-cooling, indicating the presence of bulk superconductivity. We find $T_c = 81$ K, in agreement with the values 77 K and 81 K reported for thin films^{8,12,13}.

Our preparation of pure, bulk $\text{YBa}_2\text{Cu}_4\text{O}_8$ thus permits an accurate structural determination to be made, comprising an addition to the already characterized, distinct structures of $\text{YBa}_2\text{Cu}_{3.5}\text{O}_7$ (the 247 phase; refs 5, 7) and $\text{YBa}_2\text{Cu}_3\text{O}_7$. A systematic investigation of the relationship between structure and properties in the Y-Ba-Cu-O system is therefore now possible. For example, one might address the question of why the T_c of $\text{YBa}_2\text{Cu}_{3.5}\text{O}_{7+x}$ (~40 K) is lower than that of the 124 compound; varying the stoichiometry may be revealing in this respect. We have attempted to extend the structural variations by introducing a third CuO square-planar chain between the CuO-Y-CuO layers, but so far without success.

Received 22 September; accepted 23 November 1988.

1. Zandbergen, H. W., Gronsky, R., Wang, K. & Thomas, G. *Nature* **331**, 596–599 (1988).
2. Marshall, A. F. *et al. Phys. Rev. B* **37**, 9353–9358 (1988).
3. Zandbergen, H. W. & Thomas, G. *Phys. Stat. Solidi A* **105**, 207 (1988).
4. Karpinski, J. & Kaldis, E. *Nature* **331**, 242 (1988).
5. Karpinski, J., Beeli, C., Kaldis, E., Wisard, A. & Jilek, E. *Physica C* **153**(5), 830–831 (1988).
6. Karpinski, J., Conder, K. & Kaldis, E. *Physica C* **153**(5), 401 (1988).
7. Bordet, P. *et al. Nature* **334**, 596–598 (1988).
8. Marsh, P. *et al. Nature* **334**, 141–143 (1988).
9. Fischer, P., Karpinski, J., Kaldis, E., Jilek, E. & Rusiecki, S. *Solid St. Commun.* (in the press).
10. Conder, K. & Kaldis, E. *J. less-common Metals* (in the press).
11. Müller-Bushbaum, H. K. *Angew. Chem.* **89**, 707–717 (1977).
12. Char, K., Barton, R. W., Marshall, A. F., Kapitulnik, A. & Laderman, S. S. *Physica C* **152**, 475–478 (1988).
13. Char, K., *et al. Phys. Rev. B* **38**, 834–837 (1988).

Chemical vortex dynamics in three-dimensional excitable media

Wolfgang Jahnke*, Chris Henze & Arthur T. Winfree

Department of Ecology and Evolutionary Biology,
University of Arizona, Tucson, Arizona 85721, USA

A dynamical system that can be provoked by a small stimulus to execute a large transient excursion before returning to the original state is called 'excitable'. In the Belousov-Zhabotinsky (BZ) reagent¹ such excitation propagates as a travelling pulse of oxidative activity; leaving a point source, it resembles a closed spherical surface. Activity in excitable media can also be self-organized, independent of any recent stimulus. Wavefronts in cross-section then resemble a spiral emanating from a central pivot. In three dimensions the pivot points form a line, a vortex filament, that typically closes in a slowly shrinking 'scroll ring' (Fig. 1)^{2–4}. Here we report the results of monitoring this shrinkage quantitatively both in the BZ reagent and in the Oregonator model, leading to the confirmation of mathematical arguments that derive the filament's motion from its local curvature^{5–8}. By additionally observing the predicted rapid unwinding of a coiled vortex filament⁹ we validate the reaction-diffusion theory of self-organized activity in such excitable media.

Theory suggests that a scroll ring shrinks at a rate proportional

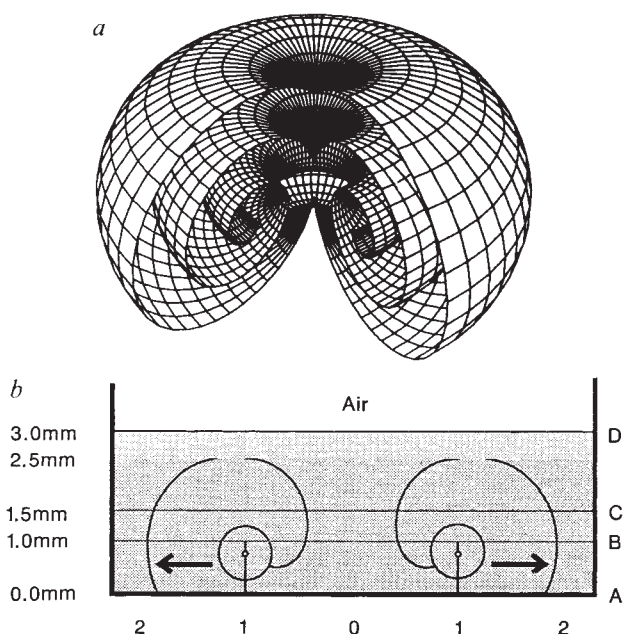


Fig. 1 *a*, Wavefronts surrounding and radiating away from a horizontal scroll ring, drawn with the front sector removed to expose interior spirals identical to those found in the more familiar two-dimensional spiral wave. For graphic clarity this is a drawing, not an actual Oregonator computation¹⁴. *b*, The scroll ring experiment is cylindrically symmetric; this diagram represents any diameter plane. Initially gel fills the dish from floor level (A) to 1.5 mm (C). Cylindrical wave 1 was started from point 0. It stays below the inexcitable layer (BC) until the upper layer (CD) is added which blocks exchange with the air. A point somewhat below the BC layer (shown as a dot) then becomes a vortex filament. The wavefront wraps into a scroll ring around it, seen here in cross-section as a pair of spirals (2). The circle is the 0.8-mm orbit of the wave tip.

to the diffusion coefficient D and inversely proportional to the local curvature or reciprocal radius R of the ring^{5–8}:

$$dR/dt = -D/R \quad (1)$$

giving

$$(R_0^2 - R^2) = 2D(t - t_0) \quad (2)$$

where R is in mm and D in $\text{mm}^2 \text{s}^{-1}$. Shrinkage proceeds until a minimum radius is reached at which the ring abruptly collapses. Because neither this minimum radius nor even the validity of the contraction rule down to small radii can be determined from theory, we have investigated the process using experiment and computation. Numerical solution of the Oregonator reaction-diffusion equations reveals a motion perpendicular to shrinkage, called 'drift', for which there is not yet a complete analytical theory^{5,6}.

In the experiments, it is essential that the filament be quantitatively observable and that the medium's properties be uniform, with no fluid convection. The chosen medium must also support spiral waves that rotate in a stable manner, without 'meander'¹⁰. Using an appropriate recipe gelled in agar at 5 °C we have met these requirements by creating a scroll ring more than 2 mm from the air interface.

For numerical simulations we use the two-variable (HBrO_2 and catalyst) Oregonator model with Tyson's 'Lo' parameters¹¹, adjusting parameters ϵ and f to match the wave speed and rotation period of vortices in our experimental medium. Computations on the Cyber 205^{12,13} use a two- or three-dimensional grid at 20- or 40- μm resolution, updated at 20-ms intervals. The inner edge of the scroll wave orbits around the vortex filament (Fig. 1); we locate this edge computationally and follow its

* Permanent address: Institut für Physikalische und Theoretische Chemie, Universität Tübingen, 7400 Tübingen, FRG.

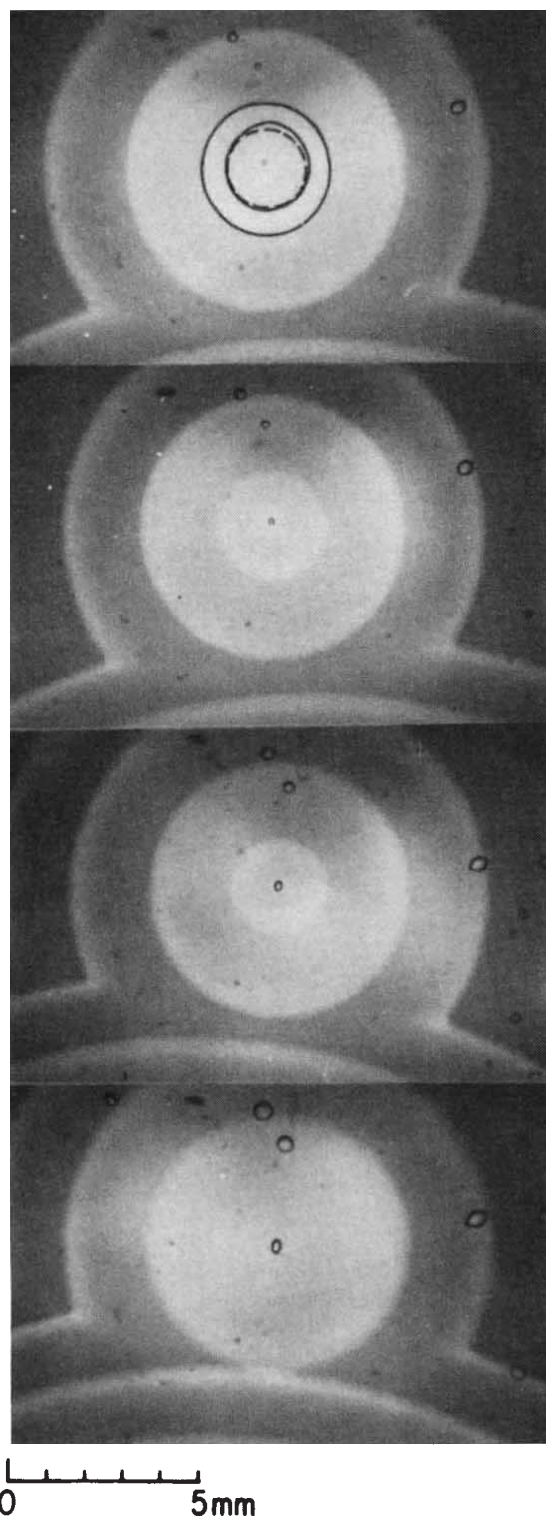


Fig. 2 The shrinkage and disappearance of a scroll ring in gelled BZ reagent (334 mM NaBrO₃, 192 mM H₂SO₄, 122 mM malonic acid, 2.9 mM ferroin in 1.3% agarose). Rotation period is 60–70 s, wavelength is 2.0 mm. Snapshots are taken 191, 253, 391 and 622 s after the second layer was added. In the first panel, the visible wavefronts are emphasized by solid curves. The centre of the vortex filament is marked by a dashed curve. Both computationally and experimentally, the vortex forms at the same depth regardless of the uniform layer's stage of recovery when poured.

motions. For both computations and experiments it is necessary to choose parameters for which rotation is simple: for recipe parameters commonly studied^{6,11}, the filament 'meanders' in tiny flower-like loops comprising two or even three distinct periodicities¹⁰.

To check equation (2) experimentally, we use a silver probe to induce a wave in a 1.5-mm layer of BZ gel, which soon grows into an expanding cylinder ~1.0 mm high. The top 0.5 mm of gel remains inexcitable because of contact with air, but becomes excitable 20 to 30 seconds after another (freshly-prepared) 1.5-mm layer is poured above it. The embedded circular wavefront edge becomes a scroll ring (Figs 1b and 2). The period and wavelength of its spirals are the same as observed in any two-dimensional spiral wave in this medium.

We have estimated the depth at which the ring forms, from experiments using a thin layer of the same gel sandwiched

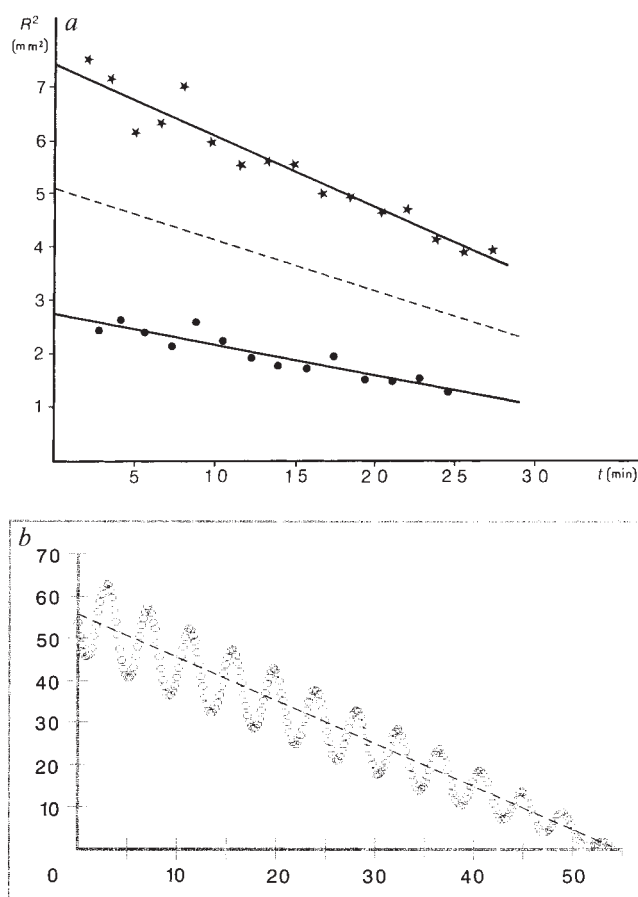


Fig. 3 *a*, Experimental data: plot of the squared radius of a scroll ring against time. This ring did not vanish within the observation time. The upper (stars) and lower (dots) rows of data correspond to the outermost and innermost positions, respectively, of the wave edge on its way around the vortex filament. The dashed middle line represents the centre of the filament. The rotation period was 95–100 s and the wavelength was 2.5 mm. Initial concentrations (before the bromination of malonic acid) were 350 mM NaBrO₃, 200 mM H₂SO₄, 122 mM malonic acid, 23 mM NaBr, 2.9 mM ferroin. *b*, Computational simulation of the experiment in *a*, using the two-variable Oregonator model with parameters $f = 1.4$, $\{\epsilon\} = 0.02$ and $D_u/D_v = 1:0.6$; these parameters are defined in refs 9–11. Data are taken at intervals of 5% of a rotation period. The maxima and minima in the curve represent the outer and inner radial positions of the wave edge in its orbit around the filament. The dashed centre line has a slope of $2D_u$ (although with other values of ϵ and f it was found to differ markedly from this theoretical expectation, sometimes even resulting in expanding scroll rings⁹. These dimensionless units of time and space scale to seconds and millimetres as functions of several chemical parameters¹⁰.

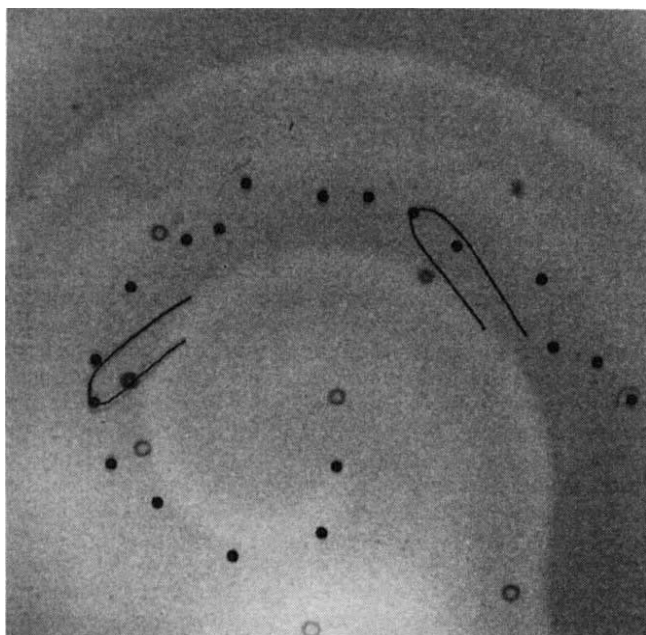


Fig. 4 A layer of well-stirred BZ reagent (same recipe as in Fig. 3), gelled above a previously gelled layer of the same reagent containing a spiral wave. The entire upper edge of the wavefront becomes a coiled horizontal scroll filament joining the pre-existing vertical filament. Because of the sharp curvature near the juncture, the filament's end retracts lengthwise and drifts laterally, straightening out its initial lateral component of curvature. Its position in time intervals of 100 s is marked by the dots. In two snapshots, the shape of the filament tip, as seen from above, is superimposed.

between glass slides. A 0.5-mm border is inexcitable until sealed against a layer of uniform gel of the same age. Any wavefront perpendicular to the edge then develops into a spiral whose tip orbits in a circle of radius 0.4 mm with centre (filament) 0.7 mm below the former surface. (This depth can be much less if the uniform gel is fresh and the buried wavefront is oblique to the surface; this angle was not independently observable in the scroll-ring experiments.)

Figure 3a shows the squared radius of a scroll ring as a function of time. Measurements are taken as the wave tip passes through the inside and outside edges of a circle of radius 0.4 mm (corresponding to wavelength/ 2π), which is the core of the spiral wave (see Fig. 1b). The dashed midline between these extremes is a measure of the contraction of the circular central filament. The squared radius of the filament decreases linearly with time from an arbitrary initial radius of 2.3 mm to a final radius of 1.6 mm after 25 min (about 16 rotations): shrinkage (dR/dt) is proportional to curvature. From equation (2) the slope of the midline ($0.0016 \text{ mm}^2 \text{ s}^{-1}$) should equal $2D$, implying $D = 0.0008 \text{ mm}^2 \text{ s}^{-1}$, which is about half of the value expected in ungelled liquid at room temperature.

To check that equation (2), stated in terms of radius of the whole circle, actually reflects a local dependence on local curvature (uniformly $1/R$ in the circular scroll ring), we also initiated elliptical rings by streaking with the silver probe. The more curved ends contract faster, as expected, until the initial eccentricity disappears.

Figure 3b corresponds to the same system as in Fig. 3a but uses Oregonator kinetics with the diffusion coefficients of the two reactants involved (HBrO_2 , ferroin) scaled in the ratio 1:0.6. Radial positions of the wave edge are plotted almost continuously as it repeatedly circles the core. As in the experiments, shrinkage is proportional to curvature, with coefficient of proportionality D in this case indistinguishable from the diffusion coefficient of the propagator variable.

While shrinking, scroll rings in BZ reagent also drift. Extinc-

tion is sometimes caused by hitting a boundary. In near-replicate experiments we have seen rings shrink to 0.5 mm in radius before hitting the boundary, upon which they commonly disappear arcwise from an initial contact point, as if making contact at a slight tilt. A vortex filament vanishes when its core (radius of 0.4 mm in these experiments) touches a wall. The ring of the system in Fig. 3a, at initial mean height $h = 0.8 \text{ mm}$, evidently did not approach the dish floor so closely while shrinking by 0.7 mm; downward drift seems slower than shrinkage. We were unable to monitor drift speed as a function of curvature but noted that the elliptical ring ruptured against the boundary first at its more-curved ends, indicating that arcs of greater curvature not only shrink faster but also drift faster.

In Oregonator computations, scroll rings drift in the direction of waves turning through the interior of the ring. In the computation represented by Fig. 3b the drift rate increases from 8% to 63% of the shrinkage rate, growing much faster than linearly with increasing curvature⁹. Scroll rings in our computations using a large array were not destroyed by hitting the boundary, but rather by shrinking and abruptly vanishing at a radius close to one core radius.

A more direct observation of the drift of scroll filaments perpendicular to their plane of curvature is obtained by an experiment suggested in ref. 9. A spiral wave in a layer of gel is overlaid with a fresh uniform layer, as described above. The upper-edge spiral of the wavefront becomes a coiled horizontal-scroll filament (Fig. 4) lying in a horizontal plane at mean height $h = 0.8 \text{ mm}$ above the dish floor. In a three-dimensional numerical experiment using Oregonator kinetics, this filament merges with the pre-existing vertical filament that extends at right angles from the dish floor, and this sharp turn then becomes more rounded¹⁴. Rounding may continue until further motion no longer changes the filaments' shape. If the curvature rule verified above in circular filaments applies also to filaments of non-uniform curvature then this shape is uniquely determined as follows. If $x(>0)$ represents the horizontal position of points along the filament and $y(x)$ represents the vertical position, an analytical solution of the appropriate differential equation⁹ yields $y(x) = a \cos^{-1}(\exp[(Dt - ax)/a^2])$, with $a \equiv 2h/\pi = 0.5 \text{ mm}$. In terms of the radius of curvature, we find $r(y) = r(0)/\cos(y/r(0))$, with $r(0) = a$. This filament arches from height h down to the dish floor and steadily slides along the floor at speed D/a , with a constant terminal radius of curvature equal to a . The endpoint thus retracts, the coil unwinding from the inside out.

The dotted arc in Fig. 4 follows the retracting endpoint of the coiled filament at intervals of 100 s, as determined from 20 successive photographs. At two points the shape of the filament tip is superimposed. The vector from each dot to the next is resolved into perpendicular components, that in the vertical plane of curvature (representing shrinkage, which component remains at $\sim 0.0050 \pm 0.0005 \text{ mm s}^{-1}$) and that in the lateral direction (representing drift, which remains $\sim 0.0028 \pm 0.0004 \text{ mm s}^{-1}$, about half of the shrinkage rate). The curvature equation (2), with coefficient D determined above by following horizontal rings to radii $> 0.5 \text{ mm}$, extrapolates to this shrinkage rate at a radius $a = 0.16 \text{ mm}$, which is less than the value of 0.5 mm predicted above. This suggests that there is a departure from equation (2) at a steadily maintained minimum radius below the limiting (core) radius of 0.4 mm estimated above. This may be due to the tightly curved tip perpetually collapsing, just as does an ephemeral scroll ring of the same radius. (Shrinkage varies nonlinearly with curvature in other excitable media, even stabilizing a vortex ring at finite radius¹⁵.) This first experimental measurement of lateral drift in a scroll ring is probably also outside the linear range.

The core of the spiral consists of chemical gradients in the filament's normal plane. Their orientation can twist with arc-length along the filament, and must do so in the form of linked rings¹⁶ or knots¹⁷. Twist is as important as curvature in directing

filaments motions^{18,19}, as is tangency to adjacent segments of filament or, equivalently, to an impermeable boundary. These factors were avoided in the discussion above, but must be reckoned with in any complete theory.

Cardiac muscle is an excitable medium in which the 'propagator species' is an electric potential and the 'catalyst' is a collection of immobile ion channel proteins; the governing partial differential equations of reaction and diffusion have the same form as for the chemically excitable medium. On the basis of this analogy, comparable vortex filaments were predicted in three-dimensional ventricular myocardium as the neuroelectric basis of re-entry in cardiac arrhythmias^{18,20,21}, and have now been found experimentally²²⁻²⁵.

A.T.W. thanks the NSF, the John D. and Catherine T. MacArthur Foundation, The Petroleum Research Foundation and the John von Neumann Supercomputer Center for support. W.J. thanks the Fulbright Commission for a fellowship to work in Tucson during 1988. We also thank P. Paluzzi of John von Neumann Supercomputer Center for the transparent volume reconstruction on the PIXAR imaging device, and Gordon Cook and Marc Courtemanche for organization and photography.

Received 30 August; accepted 4 November 1988.

1. Zaikin, A. N. & Zhabotinsky, A. M. *Nature* **225**, 535-537 (1970).
2. Winfree, A. T. *Science* **181**, 937-939 (1973).
3. Winfree, A. T. *Faraday Symp. chem. Soc.* **9**, 38-46 (1975).
4. Welsh, B., Gomatam, J. & Burgess, A. *Nature* **304**, 611-614 (1983).
5. Panfilov, A. V., Rudenko, A. V. & Krinsky, V. I. *Biofizika* **31**(5), 850-854 (1986).
6. Keener, J. P. *Physica D* **31**, 269-276 (1988).
7. Ding, D.-F. *Physica D* (in the press).
8. Keener, J. P. & Tyson, J. J. *Science* **239**, 1284-1286 (1988).
9. Winfree, A. T. & Jahnke, W. *J. phys. Chem.* (in the press).
10. Jahnke, W., Skaggs, W. E. & Winfree, A. T. *J. phys. Chem.* (in the press).
11. Keener, J. P. & Tyson, J. J. *Physica D* **21**, 307-324 (1986).
12. Nandapurkar, P. J. in *Modelling of Wave Processes in Excitable Media* (ed. Zykov, V. S.) English translation (Manchester Univ. Press, 1988).
13. Lugosi, E. & Winfree, A. T. *J. comput. Chem.* **9**(6), 689-701 (1988).
14. Henze, C., Jahnke, W., Courtemanche, M. & Winfree, A. T. in *Science at the John von Neumann Supercomputer Center Vol. 2* (Consortium for Scientific Computing, Princeton, in the press).
15. Skaggs, W. E., Lugosi, E. & Winfree, A. T. *IEEE Trans. Circ. Sys. T-CAS* **35**(7), 784-787 (1988).
16. Winfree, A. T. & Strogatz, S. H. *Physica D* **9**, 65-80 (1983).
17. Winfree, A. T. & Strogatz, S. H. *Physica D* **9**, 333-345 (1983).
18. Winfree, A. T. & Guilford, W. in *Biomathematics and Related Computational problems* (ed. Ricciardi, L. M.) 697-716 (Kluwer Academic Publishers, 1988).
19. Winfree, A. T. in *Modelling of Wave Processes in Excitable Media* (ed. Zykov, V. S.) English translation (Manchester Univ. Press, 1988).
20. Winfree, A. T. *When Time Breaks Down* 181-183 (Princeton Univ. Press, 1987).
21. Winfree, A. T. in *Science at the John von Neumann Supercomputer Center Vol. 1*, 125-130 (Consortium for Scientific Computing, Princeton, 1988).
22. Chen, P.-S. *et al. Circ. Res.* **62**, 1191-1209 (1988).
23. Shibata, N. *et al. Am. J. Physiol.* **255**, H891-H909 (1988).
24. Winfree, A. T. *J. theor. Biol.* (in the press).
25. Winfree, A. T. in *Cardiac Electrophysiology from Cell to Bedside* (eds Zipes, D. P. & Jalife, J.) (W. B. Saunders Co., in the press).

Isotopic and elemental evidence for a relationship between kimberlite and Zaire cubic diamonds

T. Akagi & A. Masuda

Laboratory for REE Microanalysis, Department of Chemistry, Faculty of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

Cubic diamonds from Zaire contain fluid inclusions which may record the conditions under which the diamonds formed—either in the mantle or in the kimberlite magma which transported them to the surface. Here we present Sr isotope ratios for whole cubic diamonds, and abundances of rare-earth elements, K, Rb, Sr and Ba for both whole cubic diamonds and the leachings from crushed cubic diamonds. The ⁸⁷Sr/⁸⁶Sr ratios of the diamonds range from 0.7038 to 0.7052, which is almost identical to that of Mbuji Mayi kimberlite¹. The chondrite-normalized elemental abundance patterns of the diamonds roughly resemble that of kimberlite but show more enrichment in incompatible elements. The bulk of the

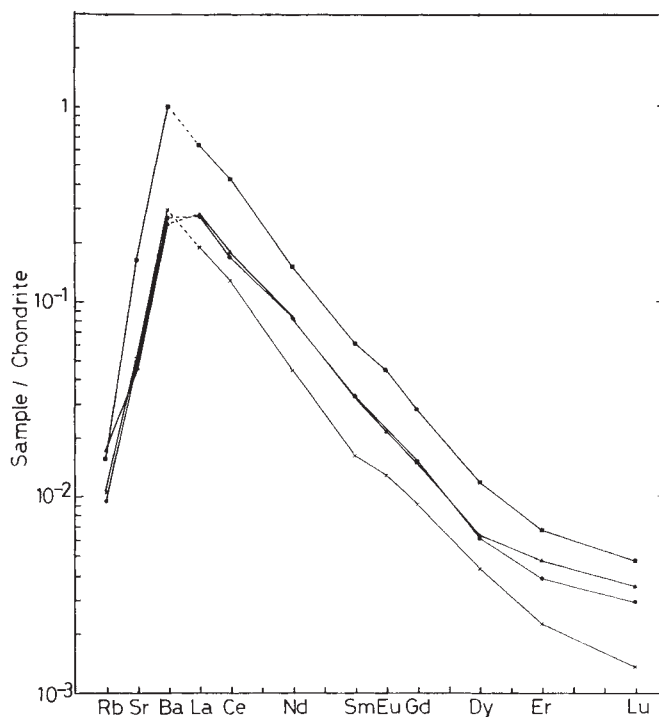


Fig. 1 Chondrite-normalized REE + (Rb, Sr, Ba) patterns of Zaire cubic diamonds. Normalized values are from refs 12 (REE + Ba), 13 (Sr) and 14 (Rb).

impurities are soluble in water. The results of the present study support the genesis of fluid inclusions in diamond in the kimberlite or in the source of kimberlite magma.

The diamonds studied here contain characteristic micro-inclusions. Previous studies on inclusions in these diamonds^{2,3} have revealed enrichment in volatiles such as water, carbon dioxide and some incompatible elements; hence, the inclusions have been thought of as fluid inclusions. Textural study of the cubic diamonds suggests strongly that the inclusions and the diamonds are syngenetic³. Meanwhile, Zashu *et al.*⁴ reported an isochron-like relation between ⁴⁰Ar and K in Zaire cubic diamonds, giving an abnormally old age of 6.0 Gyr. The normal isotopic abundance of K (refs 5, 6) rules out any explanation based on K-Ar decay; trapping of excess ⁴⁰Ar in the fluid would be the simplest explanation for the isochron-like relation⁷.

The cubic diamond samples were obtained commercially from the same source⁸ used by Zashu *et al.* They are thought to have been collected at the river beds near the Mbuji Mayi kimberlite pipe, but the identity of the diamonds' host kimberlite pipe remains uncertain. The weights of the diamonds studied were 0.2-0.5 g and the colours were green or yellow, of varying darkness. The surfaces of the diamonds were rough. Microscope studies indicated that no inclusions larger than ~10 µm in size were present near the surface of the diamonds. All diamond samples were boiled in HCl, HF, HNO₃ and HClO₄ acids before analysis.

The diamond was decomposed by combustion under flowing oxygen. The combustion apparatus used was similar to that reported previously⁵ but with two improvements. One is the covering of the inner surface of the quartz tube with a platinum vessel so that any gases or particulates released from the diamonds do not come into contact with the hot part of the quartz tube. The second improvement is the replacement of a condenser trap with a water trap for quantitative recovery. After combustion, elements on the tube and in the trap were recovered in HCl solution; this was divided into two parts, one for Sr isotope measurement and the other for abundance measurement by isotope dilution. The analytical blanks of the entire combustion method for all elements except Rb are ~1% or lower; the blank of Rb is 3-10% of the net Rb. The blank of Sr in the

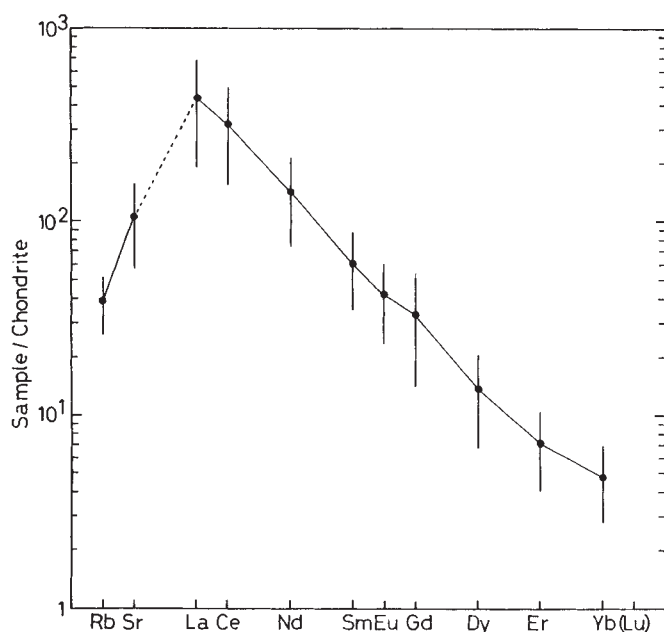


Fig. 2 Chondrite-normalized REE + (Rb, Sr) pattern of kimberlite. The concentrations are cited from ref. 8 as average kimberlite.

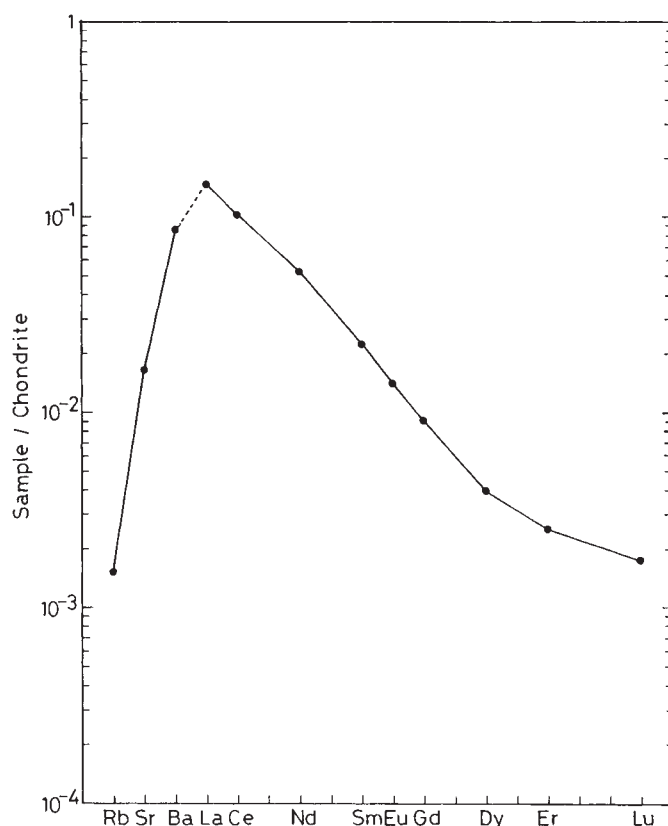


Fig. 3 Chondrite-normalized patterns of leachate (3N HCl) from crushed Zaire cubic diamonds.

combustion method is lower than 2 ng and the isotopic composition of Sr contained in the quartz used for the combustion apparatus is 0.7113. The recovered amounts of Sr by the combustion method ranged from 90 to 600 ng. The Sr ratio is not considered to suffer from significant contamination, because the Sr ratios obtained do not form a contamination plot. However, even contamination at the level of 2 ng Sr would not give any serious error in the following discussions.

Two cubic diamonds were used for the crushing experiment. A crusher made of stainless steel was used. The diamonds were crushed to the extent that no particles remained of diameter ≥ 2 mm; the crushed fragments consisted mainly of very fine, powder-like diamond particles.

The chondrite-normalized REE + (Rb, Sr, Ba) patterns of the cubic diamonds after the combustive decomposition are shown in Fig. 1. The steep slope with higher values for lighter REEs indicates enrichment in incompatible elements. Similar patterns are observed in potassium-rich rocks, such as kimberlite and lamproite. For comparison, the REE + (Rb, Sr) pattern of average kimberlite⁸ is shown in Fig. 2. The diamonds show patterns more enriched in incompatible elements. As can be seen in Fig. 1, no distinct Ce anomaly is observed, which convinces us that no significant chemical effect occurred during combustion. The slopes of the patterns in the REE region are almost constant, whereas the concentration ratios of Ba and Sr to REE vary. Taking the results of the leaching experiments into account, this diversity may correspond to a variation in the chloride/carbonate ratio in the diamond. Navon *et al.*³ have found a variation of the carbonate/water ratio in these diamonds.

The results of the leaching experiments of the crushed diamonds are shown in Fig. 3 and Table 1. The shape of the pattern of the leached HCl solution (Fig. 3) is almost the same as that of the diamond obtained by combustion, although concentration levels were about half of those found in the combustion experiment. The crushed diamonds consisted mainly of particles of size 1–100 μ m. On crushing, the diamonds presumably fracture at the inclusions. The percentages of elements leachable by sequential treatments with water, HCl, HCl again and HF are summarized in Table 1. The first treatment by water at 20–25 °C for 20 min leached out half of the K, Rb, Sr and Ba and about one tenth of REEs, and almost all remnants were leachable in HCl solution. Navon *et al.*³ observed high concentrations of carbonate and a detectable amount of chloride. Taking into account known chemical properties and solubility products, our results indicate that the water-soluble part may correspond to chloride salts, and the HCl-soluble fraction, to carbonate salts. (They are considered to have been dissolved in the fluid inclusion in the diamonds but dried after crushing.) The conclusion that the major fractions of trace elements are water-soluble provides clear evidence for the formation of fluid inclusions in this diamond.

In the combustion experiments, the elements were recovered by treatment with HCl solution; this leaching ensures that almost

Table 1 Percentages of elements leached out by sequential treatments

Order of leaching	K	Rb	Sr	Ba	La	Nd	Gd
Water	65	45	54.8	68.3	13.0	7.2	5.7
3 HCl	32	55	43.3	27.2	82.0	88.9	90.6
3 HCl	3	—	1.3	3.4	4.2	3.6	3.2
3 HF	—	—	0.6	1.1	0.8	0.3	0.5

All leaching treatments were carried out at 20–25 °C for 20 min with ultrasonic agitation.

—, Not detected.

Table 2 Rb, Sr and $^{87}\text{Sr}/^{86}\text{Sr}$ * data for the cubic diamonds

Sample	Rb (p.p.m.)	Sr (p.p.m.)	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}^\dagger$
1	0.0149	0.273	0.158	0.704030 + 43	0.70387
2	0.0289	0.887	0.0943	0.704122 + 71	0.704027
3	0.0266	0.649	0.118	0.70380 + 20	0.70368
4	0.0445	1.92	0.0671	0.70463 + 11	0.70456
5	0.0294	0.569	0.150	0.70516 + 11	0.70501

* Normalized to $^{86}\text{Sr}/^{88}\text{Sr} = 0.1194$.

† Calculated ratio for 71 Myr ago.

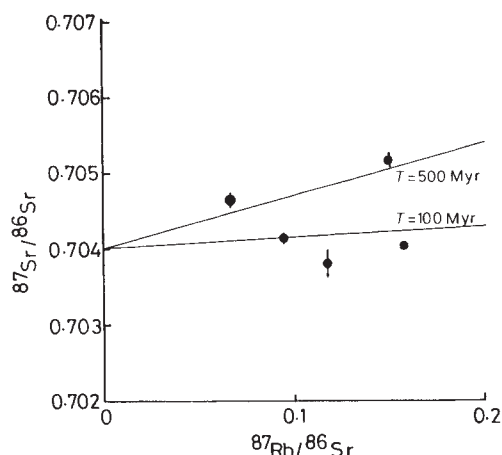


Fig. 4 $^{87}\text{Sr}/^{86}\text{Sr}$ – $^{87}\text{Rb}/^{86}\text{Sr}$ diagram for the five pieces of Zaire cubic diamond. Lines in the figures are isochrons for 100 and 500 Myr, assuming initial values of 0.704 (refs 1 and 11).

all the elements under consideration in the cubic diamonds are recovered in the combustion method.

We have measured, for the first time, isotope ratios of Sr in the cubic diamonds (Table 2). As shown in Fig. 4, the isotope ratios of the five diamonds ranged from 0.7038 to 0.7052 and the variation of the isotope ratio does not show any correlation with Rb/Sr values. We can infer that the ages of the cubic diamonds are younger than several hundred Myr (the slopes of the lines in Fig. 4 correspond to isochrons of 100 and 500 Myr) and that the diamonds formed in rather heterogeneous environments or originated from different kimberlite pipes. Alternatively, Fig. 4 might indicate the effect of high-grade metamorphism. The result also indicates that the age of the diamonds, 6.0 Gyr, obtained by an isochron-like relation in the K–Ar system, is false.

Demaiffe and Fieremans¹ reported that Sr isotope ratios of Mbuji Mayi kimberlite and Kundelunge kimberlite of age 71 Myr (eruption age of the kimberlite) are 0.7038–0.7046, being different from those of diopside xenoliths in kimberlite (0.7031 + 0.0002). Their data do not form an isochron, which indicates that the kimberlites are heterogeneous in terms of the Rb–Sr system. Barrett and Berg reported a Sr isotope ratio of ~0.704 for fresh kimberlite. Our 71-Myr Sr isotope data (see Table 2) are almost identical to the values both in kimberlite and in carbonate inclusions¹ in kimberlite.

Navon *et al.*³ consider that the fluid inclusion is pristine mantle fluid, the formation of which preceded the kimberlitic event. This conclusion is based on the difference in chemical composition between kimberlite and the diamonds. Our Sr isotope data favour a relation with kimberlite rather than with the mantle.

Boyd *et al.*⁹ showed that carbon isotope ratios of cubic diamonds have a similar $\delta^{13}\text{C}$ value to that of the coat of the octahedral diamonds, which is different from the $\delta^{13}\text{C}$ of the core. From these results, they concluded that the cubic diamonds and the coat of octahedral diamonds are formed in a different location to the core of octahedral diamonds, and that the cubic diamonds and the coat of the octahedral diamonds might have formed in the same environment, perhaps in kimberlite, which all diamonds encountered.

Our Sr isotope ratio and chemical studies of the cubic diamonds suggest that the cubic diamonds grew in the volatile-rich fluid in the kimberlite or the precursory kimberlite. The carbon isotope ratio of kimberlite shows a larger deviation than that of cubic diamonds¹⁰. To explain both the excess ^{40}Ar and the small deviation in carbon isotope ratio of the cubic diamond, some way of collecting volatiles from a large volume of the kimberlite and of averaging the carbon isotope ratio must be found⁷. Kimberlite must have had a low viscosity before eruption

and the volatiles could easily have moved in the kimberlite body. The difference in chemical compositions observed between kimberlite and the cubic diamonds (this study and ref. 3) may be mainly a result of chemical partitionings of elements between the kimberlite and the volatile-rich fluid, as incompatible elements are incorporated preferentially into the fluid.

Received 22 August; accepted 31 October 1988.

1. Demaiffe, D. & Fieremans, M. *Chem. Geol.* **31**, 311–323 (1981).
2. Melton, C. E. & Giardini, A. A. *Am. Miner.* **59**, 775–782 (1974).
3. Navon, O., Hutcheon, I. D., Rossman, G. R. & Wasserburg, G. J. *Nature* **335**, 784–789 (1988).
4. Zashu, S., Ozima, M. & Nitoh, O. *Nature* **323**, 710–712 (1986).
5. Akagi, T. & Masuda, A. *Geochem. J.* **22**, 81–82 (1988).
6. Podosek, F. A., Pier, J., Nitoh, O., Zashu, S. & Ozima, M. *Nature* **334**, 607–609 (1988).
7. Masuda, A. & Akagi, T. *Geochem. J.* **22**, 139–142 (1988).
8. Kramers, J. D., Smith, C. B., Lock, N. P., Harmon, R. S. & Boyd, F. R. *Nature* **291**, 53–56 (1981).
9. Boyd, S. R. *et al. Earth planet. Sci. Lett.* **86**, 341–353 (1987).
10. Javoy, M., Pineau, F. & Demaiffe, D. *Earth planet. Sci. Lett.* **68**, 399–412 (1984).
11. Barrett, D. R. & Berg, G. W. *Phys. Chem. Earth* **9**, 619–635 (1975).
12. Masuda, A., Nakamura, N. & Tanaka, T. *Geochim. cosmochim. Acta* **37**, 239–248 (1973).
13. Gopalan, K. & Wetherill, G. W. in *Handbook of Elemental Abundances in Meteorites* (ed. Mason, B.) 297–302 (Gordon and Breach, New York, 1971).
14. Goles, G. G. in *Handbook of Elemental Abundances in Meteorites* (ed. Mason, B.) 285–296 (Gordon and Breach, New York, 1971).

The Dupal anomaly as a trace of the upwelling lower mantle

Pat Castillo*

Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Rd, NW, Washington, DC 20015, USA

It is now widely accepted that the Earth's mantle is isotopically heterogeneous, but the scale and distribution of this heterogeneity and the structure and evolution of the mantle as a whole are poorly understood. The 'Dupal anomaly'^{1,2} has an important bearing on this problem because it is the largest isotopic domain yet delineated on the Earth's surface and it is suggested² to have been formed early in the Earth's history. Here I show that the two Dupal anomaly maxima appear to be closely associated with the two large-scale regions of low seismic velocity in the lower mantle³, which in turn are correlated with the loci of active hotspots. This correlation raises the possibility that large-scale structural features in the lower mantle produce geochemical imprints on the Earth's surface. Thus, the correlation may place severe constraints on the chemical structure of the mantle and the nature of mantle convection.

Dupré and Allègre¹ recognized that lavas from oceanic islands in the South Atlantic and Indian Ocean have high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios (>0.7035) and anomalously high $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios for a given $^{206}\text{Pb}/^{204}\text{Pb}$ ratio compared with North Atlantic and East Pacific islands. Hart² termed this isotope signature and its geographical distribution on the ocean floor the 'Dupal anomaly' and showed that although the strongest maximum of the anomaly stretches from the Indian Ocean to the South Atlantic Ocean, a second maximum with an only slightly high $^{207}\text{Pb}/^{204}\text{Pb}$ ratio exists in the southern Central Pacific Ocean. Hart² also suggested that the two maxima are connected, forming a globe-encircling belt centred on latitude 30°S. The two Dupal anomaly maxima appear to be located above the minima of the two large-scale regions of low seismic velocity (LVRs) in the lower mantle³, as well as above concentrations of the majority of active hotspots⁴ (Fig. 1).

The LVRs in the lower mantle are beneath the southern Central Pacific and the southern tip of Africa, with the latter extending beneath the Eastern Atlantic Ocean (negative contours in Fig. 1). The LVRs are surrounded by high-seismic-

* Present address: Division of Marine Geology and Geophysics, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida 33149, USA.

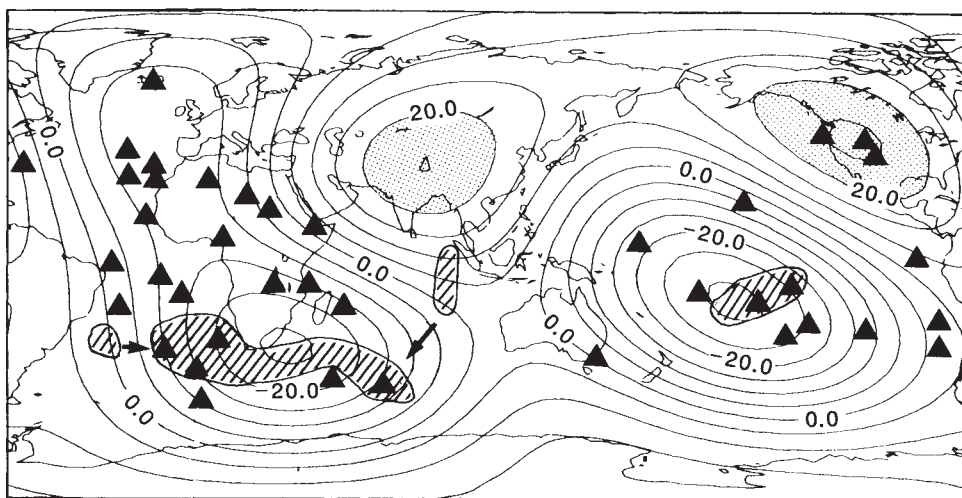


Fig. 1 World map showing model L02.56 of the three-dimensional P-wave velocity distribution averaged for the whole lower mantle using orders two and three^{3,8} (contoured in m s^{-1}), the locations of hotspots^{1,8} (triangles), and the Dupal anomaly maxima in the Indian-South Atlantic and Central Pacific regions^{1,2} (hatched areas). Dupal maxima for the Ninety-east Ridge (central hatched area) and Rio Grande Rise (far left hatched area) are samples of old hot-spot traces and should be plotted at the locations of their origin (indicated by arrows). Velocity regions higher than 20 m s^{-1} are shown stippled. Note that the Dupal anomaly maxima and the LVR minima are correlated; the majority of hotspots are above the LVRs.

velocity regions (HVRs) which are well correlated with the locations of present and geologically recent subduction zones. The good correlation between the tomographically observed HVRs and subduction zones, and more direct observations of seismically fast extensions of subduction zones deep into the lower mantle, have led some investigators⁵⁻⁸ to believe that subducted slabs continue to sink into the lower mantle, perhaps all the way to core-mantle boundary. The subducted slabs that penetrate the lower mantle must, however, be accompanied by an equal mass returning back to the upper mantle in order to conserve mass. Accordingly, Silver *et al.*⁸ suggested that LVRs represent complementary return flows of the subducted materials and that large hotspot volcanism (such as that found in Hawaii and Gough) is a manifestation of these return flows. Thus, the apparent correlation between the locations of the Dupal anomaly maxima, LVRs and hotspot concentrations suggests that the Dupal isotope signature may simply originate from hotter, upwelling regions of the lower mantle and/or the LVRs are causing at least the upper mantle, or perhaps portions of lithosphere, to produce the Dupal isotope signature. In other words, the Dupal anomaly could potentially provide a strong constraint on mantle convection, if a direct correlation between the features of the lower mantle and the geochemical properties of the Earth's surface can be established.

As a first step in establishing the correlation, it has to be shown that the two Dupal anomaly maxima, like the LVRs and hotspot concentrations, are not connected. The Dupal anomaly maxima have been delineated² by using Sr and Pb isotope data of mostly ocean island basalts (OIBs) and a few continental and subduction zone basalts. Most ocean islands, however, are built by hotspots that are generally confined within the LVRs, and data of more regional extent are therefore needed to better trace the geographical extent of the Dupal anomaly. Mid-ocean-ridge basalts (MORBs) can provide this extra needed information because unlike OIBs, which are erupted in clusters of volcanic centres, MORBs are erupted along an interconnecting, worldwide network of volcanoes. MORBs are more voluminous than OIBs and these lavas represent continuous samples of large regions of the mantle, including those regions that fill the gaps between the two Dupal anomaly maxima (that is, regions of the mantle under the East Pacific Rise (EPR) at 30°S and the Southeast Indian Ridge (SEIR)).

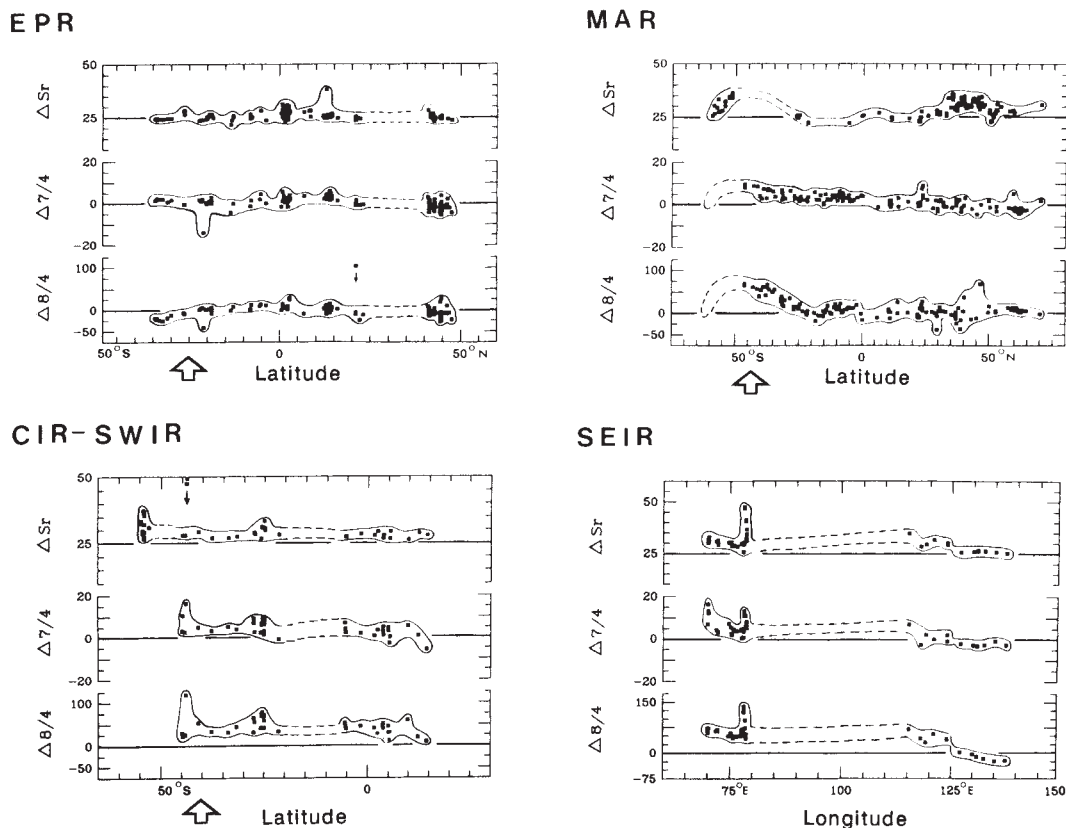
Figure 2 shows the ΔSr , $\Delta 7/4$ and $\Delta 8/4$ values of published and unpublished MORB Sr and Pb isotope data calculated according to the procedure of Hart², wherein this Δ representation is defined. North Atlantic and Pacific MORB ΔSr , $\Delta 7/4$ and $\Delta 8/4$ background values are typically ~ 25 , 0 and 0 , respectively but MORB ΔSr values that are >25 , with corresponding positive $\Delta 7/4$ and $\Delta 8/4$ values (which constitutes the Dupal

isotope signature), are present at intervals along the entire ridges. Bearing in mind the lack of data in many areas along the ridges, the Dupal anomaly appears to be strong at (1) the whole Carlsberg-Central Indian Ridge (CIR), (2) the Southwest Indian Ridge (SWIR) to $\sim 50^\circ \text{S}$ latitude, (3) the south Mid-Atlantic Ridge (MAR) from $\sim 25^\circ \text{S}$ to $\sim 55^\circ \text{S}$ latitude, and (4) the SEIR as far as $\sim 124^\circ \text{E}$ longitude¹⁰. As shown previously^{11,12}, there is no consistent and strong Dupal anomaly anywhere along the EPR. Based mainly on ΔSr and a few highly variable $\Delta 8/4$ values, there appears to be a small Dupal anomaly in Atlantic MORBs near the Azores where the extension of the African LVR is located. More Pb isotope data are needed for these lavas to confirm this point, but the work of Shirey *et al.*¹³ on Oceanographer Fracture Zone lavas at 35°N latitude, and available Sr and Pb isotope data for the Azores lavas^{2,14,15}, show clear indications of the Dupal signature.

Thus the combined data from Figs 1 and 2 show the following. (1) There is a fairly good correlation between the locations of the Dupal anomaly maxima and LVR minima in the Southern Hemisphere; both the Dupal maxima and LVR minima are disconnected and are separated by HVRs in the lower mantle³ under Australia and South America. (2) The majority of the hotspots are concentrated near LVRs; some hot spots near the LVR minima (for example, St Helena and Tubuaii), however, do not have the Dupal isotopic composition. (3) There appears to be a Dupal anomaly near the Azores, where a third but less pronounced LVR is located. (4) The magnitudes of the correlated phenomena are variable. The Dupal anomaly maximum associated with the most pronounced LVR in the Pacific generally does not have an anomalously high $\Delta 7/4$ value and it has a smaller areal extent; the anomaly maximum associated with the intermediate LVR in Africa has a larger areal extent. More recent Sr and Pb isotope data for islands and aseismic ridges near the LVR in the Pacific^{16,17} apparently do not indicate significant changes in the general areal distribution and magnitude of the Pacific anomaly. The Dupal anomaly near the Azores, if present, is the weakest in terms of the $\Delta 7/4$ and $\Delta 8/4$ values. (5) MORBs with Dupal isotopic characteristics are present at places in all ridges.

Note that the difference in isotopic compositions between MORBs from the EPR on the one hand, and MORBs from the MAR and Indian ridges (that is, those bearing the Dupal isotope signature) on the other, has been suggested^{18,19} to be due to the fast spreading rate of the EPR. Batiza¹⁸ and Allègre *et al.*¹⁹ argued that mixing of mantle components under the faster-spreading EPR is more efficient than under the slower-spreading MAR, and especially more so than under the Indian mid-ocean ridges. But Ito *et al.*²⁰ showed that this cannot be the only cause of the difference (this is also borne out in Fig. 2). ΔSr , $\Delta 7/4$

Fig. 2 ΔSr , $\Delta 7/4$ and $\Delta 8/4$ values for MORBs from the world's ridge systems plotted against locations. These values, represented by solid squares, are enclosed by solid curves (dashed curves in areas where there are no data), and background values are drawn as solid lines for clarity. A few data points (with small, solid arrows pointing downward) have unusually high values and are not enclosed within the solid curves. Errors of ΔSr , $\Delta 8/4$ and $\Delta 7/4$ values are dependent upon analytical precision of Sr and Pb isotopic ratios but are generally $\sim 2\%$, 3% and 10% respectively. Open arrows below the EPR, MAR and CIR-SWIR diagrams point to approximate locations of the Dupal anomaly belt as suggested by Hart²; the SEIR segment plotted runs sub-parallel to this belt. There is a lack of data along some segments of the ridges, but on average, values are above background at about 25° to 55° S latitude in the MAR, in the CIR-SWIR except in the northernmost part, and in the SEIR except for those east of 124° E longitude. Values for ΔSr and a few for $\Delta 8/4$ at about 30° to 50° N latitude in the MAR are also above background. The projected high ΔSr values for the south MAR are based on preliminary results of Cole *et al.*⁹. Although exhibiting some spikes, EPR values straddle the background values. Data are from Castillo, Shirey and the CHEPR team (unpublished data), Castillo and Batiza (unpublished data), and from the literature.



and $\Delta 8/4$ values between the northern and southern segments of the EPR do not differ very much, although the spreading rates of these segments vary by a factor of three. Conversely, spreading rates between the northern and southern segments of MAR do not differ significantly, although the ΔSr , $\Delta 7/4$ and $\Delta 8/4$ values of the southern segment are distinctly higher (corresponding to the Dupal signature) and more heterogeneous than those of the northern segment. The spreading rate of the SEIR east and west of the Australian-Antarctic Discordance (AAD) also does not change significantly, but the ΔSr , $\Delta 7/4$ and $\Delta 8/4$ values for MORBs west of this boundary are higher (Dupal-like) than those for MORBs to the east¹⁰.

Schilling and co-workers^{21,22} and Hamelin *et al.*²³ have suggested that the provinciality of MORB composition, particularly that of MORBs from the south MAR and Indian ridges, is due to mixing of hotspot plume(s) and the depleted component of the upper mantle. This mixing phenomenon is clearly demonstrated in MORBs near hotspots (such as Galapagos, Iceland and Bouvet). Plume influence is mostly localized, however, and is expected to decrease at a rate that is proportional to the square root of hotspot distance from the ridge²². As shown in Fig. 2, the Dupal isotope signature of MORBs covers large regions including areas well away from hotspots. For example, the 22° – 31° S segment of the MAR is supposed to be free from hotspot influence based on the plume-mixing model²², but lavas from young, isolated seamounts near this segment have the Dupal isotope signature²⁴. In comparison, lavas from young, isolated seamounts near the EPR segments remote from hotspots do not have the Dupal signature²⁴. (Lavas from young seamounts near mid-ocean ridges and MORBs come from the same upper-mantle source^{18,24–26}.) This point is also illustrated by the lavas from the SEIR segment west of the AAD, which have the Dupal isotope signature despite the fact that the segment is far

($\sim 4,000$ km) from a Dupal hotspot in the Indian Ocean. Note that the Western Victoria hotspot is near ($\sim 1,500$ km) this region of the SEIR and its lavas also have the Dupal signature². The hotspot, however, is in the east-northeast side of the AAD, wherein SEIR lavas do not have the Dupal signature. Therefore, it appears that the Dupal isotope signature of MORBs reflects the regional isotopic composition of the underlying mantle rather than the influence of hotspots. Some hotspots, especially those in the Southern Hemisphere (for example, Tristan, Gough, Kerguelen and Samoa), indeed contribute to the Dupal signature of nearby MORBs, but as a whole, hotspots have variable sizes and life spans and most importantly, their lavas have different chemical and isotopic characteristics.

In summary, there is, to a first approximation, a good correlation among the locations of the OIB Dupal anomaly maxima, LVRs and hotspot concentrations. Complementary MORB Sr and Pb isotope data strengthen this correlation. Data indicate that the Dupal anomaly can be subdivided into two (or three) sub-domains: Pacific and Indian-South Atlantic (with a minor component beneath the Azores). The locations of these Dupal maxima are correlated with the Pacific and African (plus Azores) LVR, respectively, which suggests that the Dupal isotope signature is originating from the lower mantle. Previously suggested spreading-rate and plume origins of the Dupal isotope signature in MORBs from the Indian Ocean and south MAR appear to be inconsistent with all available data.

The above correlation between tomographic features in the lower mantle and the isotopic characteristics of the oceanic crust is based on presently available data. Obviously, these are global features, and therefore more data are needed to verify or nullify such a relationship. But if the correlation exists it will give rise to questions such as why the geophysical structure of the lower mantle is reflected in the geochemistry of erupted rocks, whether

the Dupal anomaly is a tracer of motions in the lower mantle, whether the Dupal anomaly is a signature of the subcontinental mantle^{13,29,30} or of lower-mantle compositions, and whether the LVRs are return flows of previously subducted slabs. A resolution of the correlation will undoubtedly lead to further important discoveries related to chemical geodynamics^{27,28}.

The author is very grateful to P. Silver, R. Carlson, S. Shirey and J. Morris for helpful discussions. Their reviews and those by D. Gerlach, R. Batiza, J. Bender, G. Davies and friends at DTM are also acknowledged.

Received 30 May; accepted 3 November 1988.

1. Dupré, B. & Allègre, C. J. *Nature* **303**, 142–146 (1983).
2. Hart, S. R. *Nature* **309**, 753–757 (1984).
3. Dziewonski, A. M. *J. geophys. Res.* **89**, 5929–5952 (1984).
4. Stefanik, M. & Jurdy, D. *J. geophys. Res.* **89**, 9919–9925 (1984).
5. Jordan, T. H. *J. Geophys.* **43**, 473–496 (1977).
6. Craeger, K. C. & Jordan, T. H. *J. geophys. Res.* **91**, 3573–3589 (1986).
7. Silver, P. G. & Chan, W. W. *J. geophys. Res.* **91**, 13787–13802 (1986).
8. Silver, P. G., Carlson, R. W. & Olson, P. A. *Rev. Earth planet. Sci.* **16**, 477–541 (1988).

9. Cole, M. W., Hanan, B. B., Kingsley, R. & Schilling, J.-G. *Eos* **66**, 408 (1985).
10. Klein, E. M., Langmuir, C. H., Zindler, A., Staudigel, H. & Hamelin, B. *Nature* **333**, 623–629 (1988).
11. Macdougall, J. D. & Lugmair, G. W. *Earth planet. Sci. Lett.* **77**, 273–284 (1986).
12. White, W. M., Hofmann, A. W. & Purchett, H. J. *geophys. Res.* **92**, 4881–4893 (1987).
13. Shirey, S. B., Bender, J. F. & Langmuir, C. H. *Nature* **325**, 217–223 (1987).
14. Hawkesworth, C. J., Norry, M. J., Roddick, J. C. & Vollmer, R. *Nature* **280**, 28–31 (1979).
15. Dupré, B., Hamelin, B., Allègre, C. J. & Manhès, G. *US Geol. Survey Open File Rep.* **103**, 78–701 (1978).
16. Dupuy, C., Vidal, P., Barszus, H. G. & Chauvel, C. *Earth planet. Sci. Lett.* **82**, 145–152 (1987).
17. Wright, E. & White, W. M. *Earth planet. Sci. Lett.* **81**, 151–162 (1986/1987).
18. Batiza, R. *Nature* **309**, 440–441 (1984).
19. Allègre, C. J., Hamelin, B. & Dupré, B. *Earth planet. Sci. Lett.* **71**, 71–84 (1984).
20. Ito, E., White, W. M. & Gopel, C. *Chem. Geol.* **62**, 157–176 (1987).
21. Schilling, J.-H. *Earth planet. Sci. Lett.* **25**, 103–115 (1975).
22. Hanan, B. B., Kingsley, R. H. & Schilling, J.-G. *Nature* **322**, 137–144 (1986).
23. Hamelin, B., Dupré, B. & Allègre, C. J. *Earth planet. Sci. Lett.* **67**, 340–350 (1984).
24. Batiza, R. & Castillo, P. R. *Eos* **69**, 491 (1988).
25. Zindler, A., Staudigel, H. & Batiza, R. *Earth planet. Sci. Lett.* **70**, 175–195 (1984).
26. Batiza, R. & Vanko, D. *J. geophys. Res.* **89**, 11235–11260 (1984).
27. Allègre, C. J. *Tectonophysics* **81**, 395–400 (1982).
28. Zindler, A. & Hart, S. R. A. *Rev. Earth planet. Sci.* **14**, 493–571 (1986).
29. Gerlach, D. G., Hart, S. R., Morales, V. W. & Palacios, C. *Nature* **322**, 165–169 (1986).
30. Hawkesworth, C. J., Mantovani, M. S. M., Taylor, P. N. & Palacs, Z. *Nature* **322**, 356–359 (1986).

Theoretical evidence for a new ultra-high-pressure phase of SiO₂

Key Taek Park*, Kiyoyuki Terakura* & Yoshito Matsui†

* Institute for Solid State Physics, University of Tokyo, Roppongi, Minato-ku, Tokyo 106, Japan

† Institute for Study of the Earth's Interior, Okayama University, Misasa, Tottori-ken 682-02, Japan

A phase of silica denser than stishovite (SiO₂-rutile) has been searched for intensively. Following the suggestions by Al'tshuler *et al.*¹ and Simakov *et al.*², SiO₂-fluorite was once regarded as a candidate for such a high-pressure phase. Subsequent theoretical studies^{3,4} revealed, however, that SiO₂-fluorite would have at most about the same density as SiO₂-rutile and that it would be dynamically unstable at pressures below 170 GPa (ref. 4). Here we propose an alternative hypothetical polymorph of silica with a modified fluorite (Pa $\bar{3}$) structure as a possible high-pressure phase. SiO₂-Pa $\bar{3}$ would have a density of 4.46 g cm⁻³ (~6% denser than SiO₂-rutile at normal pressures) and should become more stable than SiO₂-rutile at pressures of ≥ 60 GPa. These results also suggest that MgSiO₃-perovskite, which is widely accepted as the major constituent within the Earth's deep interior, may be unstable at very high pressure.

The present arguments are based on first-principles electronic-structure calculations for SiO₂-fluorite, -rutile and -Pa $\bar{3}$, the crystal structures of which are shown in Fig. 1. SiO₂-Pa $\bar{3}$ was predicted theoretically as a stable phase in a molecular-dynamics study of SiO₂-fluorite⁵. However, its optimal structure and relative stability with respect to other phases could not be inferred unambiguously from this previous work because of the use of simplified classical interatomic potentials with uncertain ranges of applicability. We have attempted to obtain such information using first-principles electronic-structure calculations.

We used the local-density approximation (LDA) in density functional theory⁶ and the full-potential linear augmented plane wave (FLAPW) method^{7,8} to solve the resultant effective one-electron problem. To limit the number of plane-wave basis functions, we used an energy cutoff of 22 Rydbergs, which results in a maximum dimension of ~850 for the hamiltonian matrix of SiO₂-Pa $\bar{3}$. The numbers of *k* points in the first Brillouin zone for the *k*-space integration are 32, 8 (16) and 8 (32) for SiO₂-fluorite, -rutile and -Pa $\bar{3}$, respectively. (The numbers in the parentheses denote the numbers of *k* points equivalent to SiO₂-fluorite for equal numbers of atoms in a unit cell.) For insulators, these numbers are sufficient to obtain structural information. We found that the results for SiO₂-rutile were affected only marginally by doubling the number of *k* points used. The 3s and 3p states of Si and the 2s and 2p states of O were treated as valence states and we adopted a single energy window, which is sufficient when, as in this case, there is no problem of shallow semicores. The combination of the LDA and the pseudo-potential method has been applied successfully to several systems, including α -cristobalite⁹. As the FLAPW method is no less accurate than the pseudo-potential method, the present approach is expected to be fairly reliable for discussing the structure of silica. The reliability of our calculations may be judged by Table 1, where our results are compared with those of other theoretical studies and also with available experimental data.

Figure 2 shows the calculated total energies as a function of volume for SiO₂-fluorite, -rutile and -Pa $\bar{3}$. For SiO₂-fluorite, the crystal structure is uniquely determined by its volume, and the total-energy calculation for a given volume is straightforward. The present result, denoted by the long-dashed curve, agrees fairly well with those from other sources, as shown in Table 1. On the other hand, for SiO₂-rutile and -Pa $\bar{3}$, arbitrary parameters (*c/a* and *u* for rutile and *u* for Pa $\bar{3}$) arise in the optimization of the structure for a given volume. Here, *a* and *c* are unit-cell dimensions; for SiO₂-rutile, *u* is defined such that the nearest Si-O distance in the *c*-plane is given by $\sqrt{2}ua$, whereas in

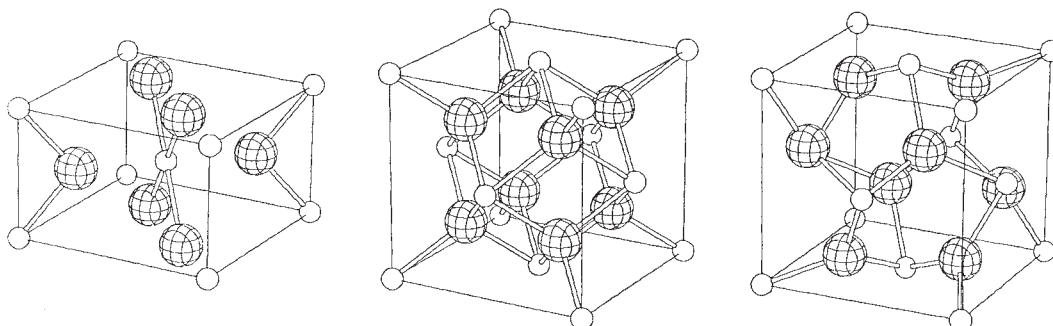


Fig. 1 From left to right, the unit cells of SiO₂ with rutile, fluorite and Pa $\bar{3}$ structures, respectively. Si is denoted by a small open sphere and O by a large hatched one. Nearest-neighbour atoms are connected by bars.

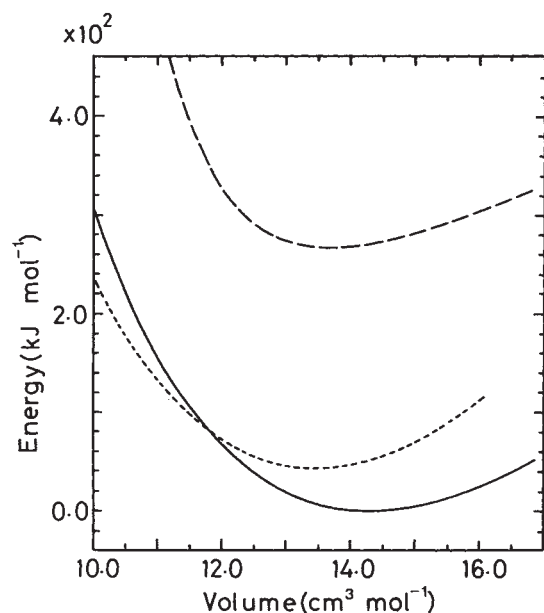


Fig. 2 The calculated total energies as a function of volume for the three crystal structures of SiO_2 . The long-dashed line denotes the fluorite, the short-dashed line denotes the $\text{Pa}\bar{3}$ and the solid line denotes the rutile structures. The energy of SiO_2 -rutile at zero pressure is taken as the origin of energy.

SiO_2 - $\text{Pa}\bar{3}$, u defines one of the eight c sites for the oxygens, which is given by $(u, u, u)a$ (ref. 10).

In most of the calculations for SiO_2 -rutile and $-\text{Pa}\bar{3}$, the structural parameters were fixed either at the observed values (for rutile)^{11,12} or at the optimized values (for $\text{Pa}\bar{3}$) at normal pressures. For some volumes, however, the total energies were calculated for the fully optimized structures. The results for the approximately optimal structures were corrected by extrapolation or interpolation on the basis of the fully optimized ones. The corrected data were then fitted with the Murnaghan equation of state¹³, as denoted in Fig. 2 by the solid curve for rutile and the short-dashed curve for $\text{Pa}\bar{3}$. The optimal value of u for SiO_2 -rutile in our calculations ($u=0.303$) agrees with the observed value¹¹ within 0.1% error. For SiO_2 - $\text{Pa}\bar{3}$, our u value at normal pressure is 0.344, whereas semiempirical analysis⁵ gives 0.3349. For the optimized SiO_2 - $\text{Pa}\bar{3}$ structure, the scatter in the neighbouring O-O distances is smaller than that in SiO_2 -rutile, and the Si-O interatomic distance is shorter than the O-O distance; the latter contrasts with FeS_2 - $\text{Pa}\bar{3}$ (pyrite), in which $u=0.386$ and the S-S distance indicates a disulphide ion¹⁴. Note also that the fluorite structure is a special case of the $\text{Pa}\bar{3}$ structure for which $u=0.25$. For these reasons we refer to SiO_2 - $\text{Pa}\bar{3}$ as a modified fluorite rather than as a pyrite¹⁵.

In Figs 3 and 4 we show, respectively, the enthalpy-pressure ($H(p)$) and density-pressure ($\rho(p)$) relations derived from the results in Fig. 2. Several important conclusions can be drawn from these figures. First, the SiO_2 -fluorite phase is energetically unfavourable in the entire pressure range. The compressibility of SiO_2 -fluorite decreases rapidly as pressure increases. Second, SiO_2 -rutile is the most stable of the three structures at normal pressures, as we should expect. The calculated equilibrium density is lower than the observed one by 2%, which causes a reduction in the bulk modulus (Table 1). Third, our central result is that SiO_2 - $\text{Pa}\bar{3}$ is energetically more stable than SiO_2 -rutile at pressures ≥ 60 GPa. Even at normal pressures there are indications that this is the case: the density of SiO_2 - $\text{Pa}\bar{3}$ is higher, and the pressure derivative of bulk modulus, dB/dp , is significantly smaller, than those of SiO_2 -rutile. It should be noted that the results in Figs 2-4 are all for a static lattice at zero temperature. The transition pressure would not be significantly different at finite temperature, however, because it is probable

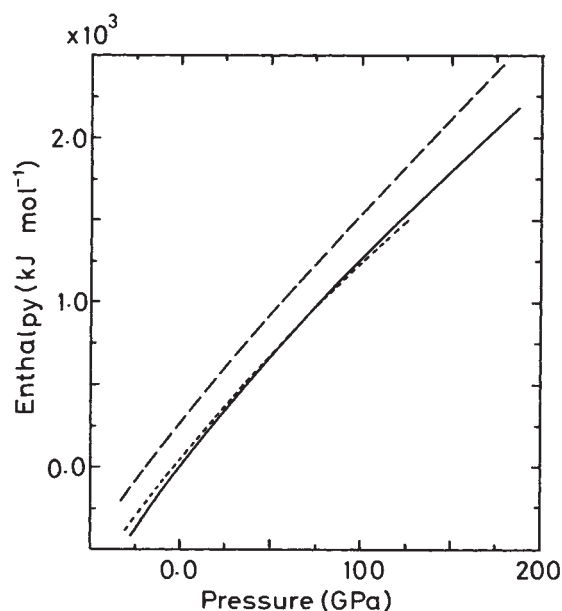


Fig. 3 Calculated enthalpy against pressure for three crystal structures of SiO_2 . The key is the same as in Fig. 2.

that the entropy change on transitions between crystalline polymorphs is generally very small.

One must ask why SiO_2 - $\text{Pa}\bar{3}$ has not been observed experimentally at pressures >100 GPa. There are several possible reasons for this. First, there is a possibility that the transition pressure is underestimated in our calculations. Figure 3 suggests that a small error in the total-energy calculation is amplified in estimating the transition pressure. This was indeed the case in a similar calculation for polymorphs of Si (ref. 16). Second, SiO_2 -rutile undergoes a structural phase transition to a CaCl_2 structure below 60 GPa (T. Yagi, private communication), involving a slight reduction in the enthalpy, and the transition pressure to SiO_2 - $\text{Pa}\bar{3}$ will thereby be shifted to a slightly higher value. Third, there is a significant rearrangement of atoms associated with the structural transformation from the rutile-type or CaCl_2 -type to the $\text{Pa}\bar{3}$ -type structure, so that the activation energy may be too large for such a transformation to be realized under normal experimental conditions.

Finally, we point out some important implications of this work. Available experimental and theoretical data¹⁷⁻¹⁹ together with our results indicate that the density of the assemblage of SiO - $\text{Pa}\bar{3}$ plus MgO in the sodium chloride structure is lower than that of MgSiO_3 -perovskite by only 1% at normal pressures and becomes higher by $\sim 4\%$ at 100 GPa. Such a cross-over cannot be expected at least up to 150 GPa for SiO_2 -rutile. Our results suggest that MgSiO_3 -perovskite might be less stable than the assemblage of SiO_2 - $\text{Pa}\bar{3}$ plus MgO-(sodium chloride) at very

Table 1 Calculated and experimental properties of SiO_2

	Bulk modulus B (GPa)	dB/dp	Equilibrium density ρ_0 (g cm^{-3})	Total energy* $E_{\min}$ (kJ mol^{-1})
Fluorite†	328	10.73	4.39	267
Fluorite‡	330		4.30	
Fluorite§	310		4.30	
$\text{Pa}\bar{3}$ †	335	1.60	4.46	43
Rutile†	288	3.14	4.20	0
Rutile	335	5.70	4.28	

* Relative to rutile.

† Present FLAPW calculation.

‡ ASW calculation³.

§ APW calculation⁴.

|| Experiment¹².

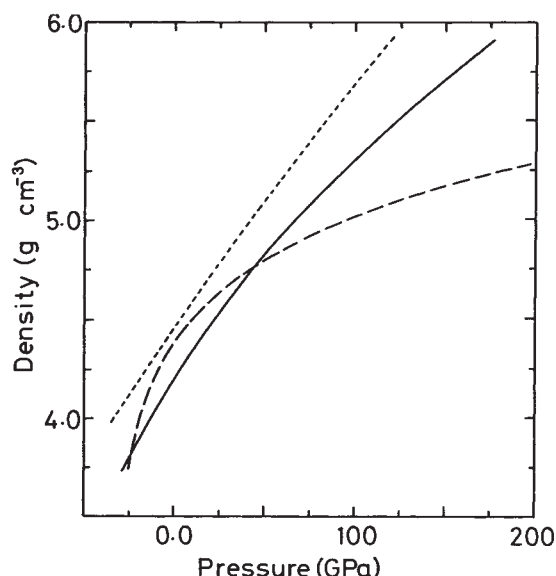


Fig. 4 Calculated density against pressure for three crystal structures of SiO_2 . Same key as for Fig. 2.

high pressures. At present, the pressure required for such a transition cannot be estimated because of the lack of reliable information on the differences in internal energies. A high-temperature diamond-cell experiment¹⁸ shows that $(\text{Mg}_{0.88}\text{Fe}_{0.12})\text{SiO}_3$ -perovskite is stable up to at least 120 GPa, but this does not completely exclude the possibility of a transition pressure lower than 120 GPa, because of the question of activation energy mentioned above. To model realistically conditions in the Earth's lower mantle, one must include the effect of iron. It is pointed out that iron in FeO (where it has sixfold coordination) would undergo a high-spin to low-spin transition in the lower mantle, accompanied by an appreciable volume contraction, whereas iron in $(\text{Mg}, \text{Fe})\text{SiO}_3$ (eightfold coordination) would remain in a high-spin state under similar conditions²⁰. The presence of Fe may therefore encourage the sort of transition we are proposing. We plan a more extensive study on silicate-perovskite. Our identification of SiO_2 - $\text{Pa}\bar{3}$ as one of the high-pressure phases of silica provides motivation for a search for other such phases. Although SiO_2 - $\text{Pa}\bar{3}$ contains 12 atoms per unit cell, the structure is rather simple because of its cubic symmetry. SiO_2 - $\text{Pa}\bar{3}$ may correspond to a local energy minimum in a rather narrow region of the phase space describing the polymorphs of silica, but the global minimum may prove harder to identify.

The authors thank Professor T. Yagi for providing them with unpublished data for the pressure-induced structural transformation of SiO_2 -rutile. K.T. is grateful to Dr N. Hamada and Professor A. Yanase for their collaboration in constructing the FLAPW program package. This work is supported in part by a Grant-in-Aid for Scientific Research on Priority Areas by the Ministry of Education, Science and Culture. Numerical computations were performed at the Computer Centres of the Institute for Solid State Physics, the Institute for Molecular Science, the University of Tokyo and Kyoto University.

Received 19 August; accepted 7 November 1988.

1. Al'tshuler, L. V., Podurets, M. A., Simakov, G. V. & Trunin, R. F. *Soviet Phys. Solid St.* (Engl. Trans.) **15**, 969-971 (1973).
2. Simakov, G. V., Podurets, M. S. & Trunin, R. F. *Dokl. Akad. Nauk SSSR* **211**, 1330-1332 (1973).
3. Carlsson, A. E., Ashcroft, N. W. & Williams, A. R. *Geophys. Res. Lett.* **11**, 617-619 (1984).
4. Bukowinski, M. S. T. & Wolf, G. H. *J. geophys. Res.* **91**, 4704-4710 (1986).
5. Matsui, Y. & Matsui, M. in *Structural and Magnetic Phase Transitions in Minerals, Advances in Physical Geochemistry* Vol. 7 (eds Ghose, S., Salje, E. & Coey, J. M. D.) 129-140 (Springer, New York, 1988).
6. Kohn, W. & Sham, L. J. *Phys. Rev.* **A140**, 1133-1138 (1965).
7. Weinert, E., Wimmer, W. & Freeman, A. J. *Phys. Rev.* **B26**, 4571-4578 (1982).
8. Park, K. T., Terakura, K., Oguchi, T., Yanase, A. & Ikeda, M. *J. phys. Soc. Japan* **57**, 3445-3456 (1988).

9. Allan, D. C. & Teter, M. P. *Phys. Rev. Lett.* **59**, 1136-1139 (1987).
10. Hahn, T. *International Tables for Crystallography* Vol. A, 620-621 (Reidel, Boston, 1983).
11. Baur, W. H. & Khan, A. A. *Acta Crystallogr.* **B27**, 2133-2139 (1971).
12. Liu, L., Bassett, W. A. & Takahashi, T. *J. geophys. Res.* **79**, 1160-1164 (1974).
13. Murnaghan, F. D. *Proc. natn. Acad. Sci. U.S.A.* **30**, 244-247 (1944).
14. Wyckoff, R. W. G. *Crystal Structures* 2nd edn. Vol. 1, 346 (Wiley, New York, 1948).
15. Matsui, Y. & Kawamura, K. in *High-Pressure Research in Mineral Physics* (eds Manghnani, M. H. & Syono, Y.) 305-311 (Terra Scientific/American Geophysical Union, Tokyo/Washington, DC, 1987).
16. Yin, M. T. & Cohen, M. L. *Phys. Rev.* **B26**, 5668-5687 (1982).
17. Bukowinski, M. S. T. *Geophys. Res. Lett.* **12**, 536-539 (1986).
18. Knittle, E. & Jeanloz, R. *Science* **235**, 668-670 (1987).
19. Bukowinski, M. S. T. & Wolf, G. H. in *Structural and Magnetic Phase Transitions in Minerals, Advances in Physical Geochemistry* Vol. 7 (eds Ghose, S., Salje, E. & Coey, J. M. D.) 91-112 (Springer, New York, 1988).
20. Sherman, D. M. in *Structural and Magnetic Phase Transitions in Minerals, Advances in Physical Geochemistry* Vol. 7 (eds Ghose, S., Salje, E. & Coey, J. M. D.) 113-128 (Springer, New York, 1988).

Development of neuronal polarity: GAP-43 distinguishes axonal from dendritic growth cones

Kimberly Goslin*, David J. Schreyer†, J. H. Pate Skene† & Gary Banker*

* Department of Anatomy, Cell Biology and Neurobiology, Albany Medical College, Albany, New York 12208, USA

† Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

Outgrowth of distinct axonal and dendritic processes is essential for the development of the functional polarity of nerve cells. In cultures of neurons from the hippocampus, where the differential outgrowth of axons and dendrites is readily discernible, we have sought molecules that might underlie the distinct modes of elongation of these two types of processes. One particularly interesting protein is GAP-43 (also termed B-50, F1 or P-57), a neuron-specific, membrane-associated phosphoprotein whose expression is dramatically elevated during neuronal development and regeneration¹⁻⁷. GAP-43 is among the most abundant proteins in neuronal growth cones^{8,9}, the motile structures that form the tips of advancing neurites, but its function in neuronal growth remains unknown. Using immunofluorescence staining, we show that GAP-43 is present in axons and concentrated in axonal growth cones of hippocampal neurons in culture. Surprisingly, we could not detect GAP-43 in growing dendrites and dendritic growth cones. These results show that GAP-43 is compartmentalized in developing nerve cells and provide the first direct evidence of important molecular differences between axonal and dendritic growth cones. The sorting and selective transport of GAP-43 may give axons and axonal growth cones certain of their distinctive properties, such as the ability to grow rapidly over long distances or the manner in which they recognize and respond to cues in their environment.

The axons and dendrites that develop in hippocampal cultures differ from one another in form, in molecular composition, and in synaptic polarity, just as their counterparts do *in situ*¹⁰⁻¹². They also differ in their growth characteristics. Axons, which arise after one day in culture, are long, thin, rapidly elongating processes of relatively uniform diameter¹¹. Dendrites begin to elongate a few days later and are short, highly tapered, and slow-growing. With continued development, the cells become interconnected by a dense network of axonal processes, which run along the cell bodies and dendrites forming numerous synaptic contacts¹².

We examined the distribution of GAP-43 immunoreactivity in cultured hippocampal neurons, using fluorescence microscopy. Figure 1 illustrates GAP-43 immunostaining of hippocampal neurons after six days in culture, when both axons and dendrites are actively growing¹¹. At this stage of development dendrites and axons can be readily distinguished by their morphology. GAP-43 immunoreactivity was detected in the somata of virtually all neurons in these cultures. In addition,

Fig. 1 A comparison of GAP-43 immunoreactivity in axonal and dendritic growth cones. The phase contrast (*a*, *c*, *e*) and corresponding fluorescence photomicrographs (*b*, *d*, *f*) illustrate the distribution of GAP-43 in hippocampal neurons after six days in culture, as revealed by immunostaining with the monoclonal antibody 91E12. *a*, *b*, Two axons, one of which ends in a prominent growth cone (arrowhead). GAP-43 immunoreactivity is present throughout the axons and is highly concentrated in the growth cone. *c*, *e*, Two cells, their dendrites, and dendritic growth cones (arrows). Several axons that arise from neighbouring cells traverse each field and run along some of the dendrites. In the corresponding immunofluorescence micrographs, dendritic growth cones are unstained (arrows). Dendrites, despite their large diameter, are also not visible by immunofluorescence. In contrast, even axons of fine calibre, barely detectable in the phase contrast micrographs, are brightly stained for GAP-43 (arrowheads in *c* and *d*). In *e* and *f*, a single axon (arrowheads) enters the field from below and runs along the cell body and one of the dendrites. Just before its termination, the axon diverges from the dendrite, so that their growth cones are adjacent. The marked difference in their staining is apparent. Scale bar, 15 μ m.

Methods. Hippocampal cells were obtained from 18-day-old embryonic rat brains, plated onto polylysine-treated glass coverslips, and maintained in serum-free medium, as described previously^{29,30}. They were fixed in 4% formaldehyde in phosphate-buffered saline containing 0.12 M sucrose, permeabilized in ethanol, then incubated with 91E12 (1:4,000), followed by fluorescein-labelled goat anti-mouse IgG (1:400). The antibody recognizes a single band on immunoblots of proteins from hippocampal cultures (data not shown). It reacts with both newly synthesized and post-translationally modified forms of the protein (D.S. and P.S., unpublished observations).

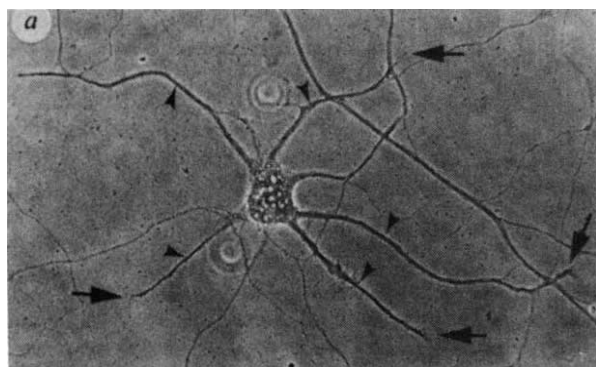
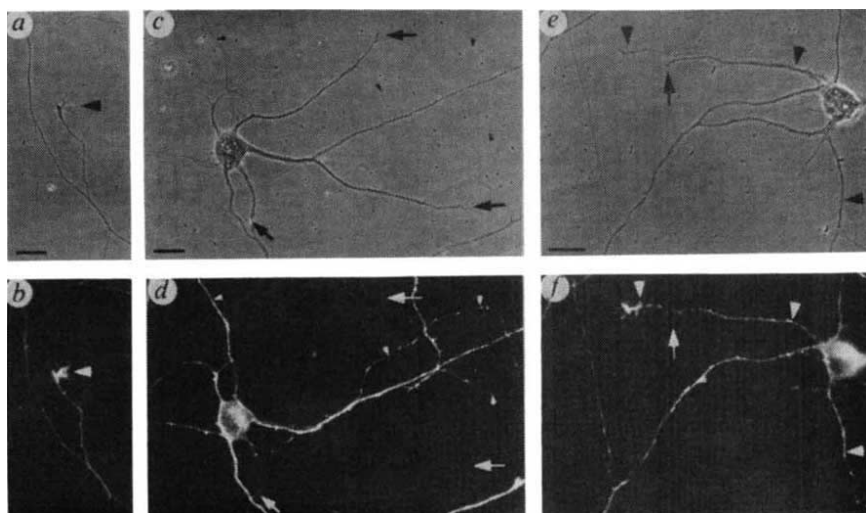
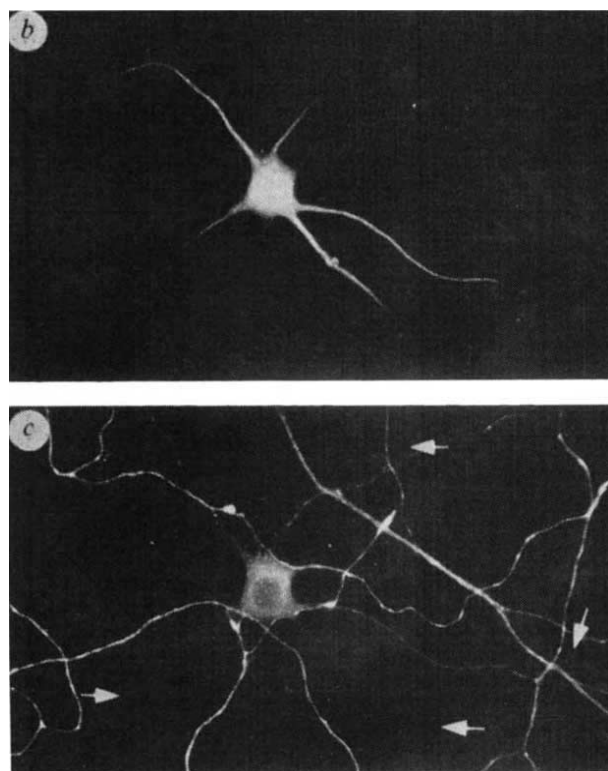


Fig. 2 Simultaneous localization of MAP2 and GAP-43 in an eight-day-old hippocampal neuron. Processes identified as dendrites by the presence of MAP2 completely lack GAP-43 immunoreactivity. *a*, Phase-contrast micrograph showing the five dendrites arising from the cell (arrowheads), four of which terminate in obvious growth cones (arrows). *b*, MAP2 immunostaining is present throughout the dendrites, but ends just proximal to their growth cones. *c*, The dendrites and their growth cones are not visible after immunostaining for GAP-43. The many axons that traverse the field, as seen in the phase-contrast micrograph and by GAP-43 immunofluorescence, are not stained with MAP2. Scale bar, 15 μ m.

Methods. Cultures, prepared and fixed as described in Fig. 1, were reacted simultaneously with a polyclonal antibody against GAP-43 and a monoclonal antibody, AP14, against MAP2, both at dilutions of 1:200. GAP-43 was visualized with fluorescein-labelled goat anti-rabbit IgG and MAP2 with biotinylated horse anti-mouse IgG followed by rhodamine-labelled avidin. Pre-immune serum gave only faint staining of cell bodies (not shown).



immunofluorescence was present at high levels throughout the axonal processes, and was concentrated in axonal growth cones, as previously described^{9,13}. In sharp contrast, dendritic growth cones had no GAP-43 immunoreactivity, even though in cultures of this age, dendritic growth cones, like axonal growth cones, are highly motile and actively elongating (K.G. and G.B., unpublished observations). The shafts of dendritic processes were also unstained, but axonal processes running along the dendrites were readily visible by their bright immunofluorescence. The flat, non-neuronal cells in these cultures lacked GAP-43 immunoreactivity (data not shown).

To verify that GAP-43 was associated exclusively with axonal processes, we used double-label immunofluorescence to localize simultaneously GAP-43 and the microtubule-associated protein MAP2 (Fig. 2). Previous work has shown that MAP2 is restricted to cell bodies and dendrites in hippocampal cultures¹⁴. Both GAP-43 and MAP2 were present in neuronal cell bodies, but their distribution in processes was complementary; GAP-43 was completely absent from processes that contained MAP2, and MAP2 was completely absent from processes that contained GAP-43. Virtually every neurite observed contained either one or the other antigen. MAP2 immunoreactivity did not extend

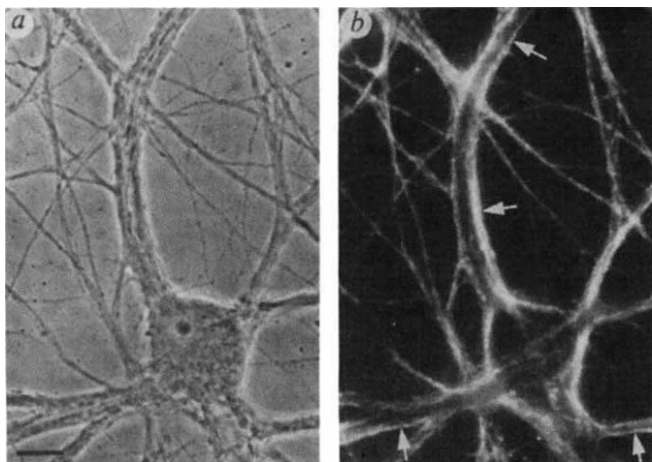


Fig. 3 GAP-43 immunoreactivity in a mature hippocampal neuron. Phase contrast (*a*) and immunofluorescence micrographs (*b*) of a cell fixed after 28 days in culture. This cell gives rise to several main dendrites. Axons running along the dendrites are stained for GAP-43, but the dendrites themselves are unstained. This is most clearly seen where axons lie along the edges of the dendrites, but do not run over the dendritic surface (arrows in *b*). Methods as in Fig. 1. Scale bar, 10 μ m.

into the dendritic growth cones, which were nevertheless clearly visible in the corresponding phase contrast photomicrographs; they were completely unstained with antiserum against GAP-43. The pattern of staining observed with both a polyclonal antiserum against GAP-43 and a monoclonal antibody that reacts with newly synthesized and post-translationally modified GAP-43 (D.S. and P.S., unpublished observations) was identical. It therefore seems unlikely that GAP-43 was present in dendrites but not detected by these antibodies.

In addition to its widespread association with growing axons, GAP-43 remains abundant in a subset of neurons in adult brain¹⁵. We therefore examined the distribution of GAP-43 in more mature hippocampal cultures, where synaptic connections are numerous¹². A neuron from such a culture, maintained for 28 days, is shown in Fig. 3. As at earlier stages, GAP-43 immunoreactivity was strictly confined to the axonal domain. Dendrites seemed strikingly empty in comparison to the brightly stained axons that ran along their surface. The polarized distribution of GAP-43 in these cultures suggests that, at mature synapses, GAP-43 is exclusively presynaptic (see also ref. 16).

These observations have important implications for the function of GAP-43 during neuronal growth. Its presence in growth cones cannot be an absolute requirement for neurite extension in general: even when dendrites are actively elongating, GAP-43 cannot be detected in their growth cones. Instead, GAP-43 may be important for aspects of growth that are specific to axons. For example, axons are distinguished by their rapid elongation velocity, which requires rapid membrane addition and cytoskeletal assembly, processes thought to occur at the growth cone where GAP-43 is concentrated¹⁷⁻¹⁹. In addition, GAP-43 might modulate the response of axonal growth cones to environmental cues important in pathfinding or target recognition^{20,21}. GAP-43 binds calmodulin in the absence of free calcium²², and inhibits production of phosphatidyl-4,5-bisphosphate, the immediate precursor for the calcium-mobilizing second messenger inositol 1,4,5-trisphosphate^{23,24}. These actions are consistent with the involvement of GAP-43 in the modulation of intracellular signalling mediated by calcium, an important regulator of growth cone function^{25,26}. The distinct growth characteristics of dendrites raise the possibility that their growth is regulated by different mechanisms. Our results, which provide the first direct evidence of molecular differences between axonal and dendritic growth cones, are consistent with this idea.

The distribution of GAP-43 observed in hippocampal neurons implies that, after its synthesis in the cell body, it must be preferentially routed into the axon. *In vivo*, GAP-43 is one of a group of rapidly transported proteins that are believed to move along axons in association with assembled membrane vesicles^{27,28}. This differential routing of GAP-43 into axons may, therefore, reflect a more general selection of different membrane populations for transport into growing axons and dendrites, a process of fundamental importance to the development of neuronal polarity.

This research was supported by the NIH. The antibody against MAP2 was a gift from L. Binder.

Received 14 September; accepted 18 October 1988.

- Skene, J. H. P. *A. Rev. Neurosci.* (in the press).
- Benowitz, L. I. & Routenberg, A. *Trends Neurosci.* **10**, 527-532 (1987).
- Skene, J. H. P. & Willard, M. *J. Cell Biol.* **89**, 86-95 (1981).
- Skene, J. H. P. & Willard, M. *J. Cell Biol.* **89**, 96-101 (1981).
- Benowitz, L. I. & Lewis, E. R. *J. Neurosci.* **3**, 2153-2163 (1983).
- Redshaw, J. D. & Bixby, M. A. *Can. J. Physiol. Pharmac.* **62**, 1387-1393 (1984).
- Jacobson, R. D., Virag, I. & Skene, J. H. P. *J. Neurosci.* **6**, 1843-1855 (1986).
- Skene, J. H. P. *et al. Science* **233**, 783-786 (1986).
- Meiri, K. F., Pfenninger, K. & Willard, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3537-3541 (1986).
- Banker, G. A. & Waxman, A. in *Intrinsic Determinants of Neuronal Form and Function* (eds R. Lasek & M. B. Black.) 61-82 (Liss, New York, 1988).
- Dotti, C. G., Sullivan, C. A. & Banker, G. A. *J. Neurosci.* **8**, 1454-1468 (1988).
- Bartlett, W. P. & Banker, G. A. *J. Neurosci.* **4**, 1954-1965 (1984).
- Meiri, K. F., Willard, M. & Johnson, M. I. *J. Neurosci.* **8**, 2571-2581 (1988).
- Caceres, A., Banker, G. & Binder, L. *J. Neurosci.* **6**, 714-722 (1986).
- Benowitz, L. I., Apostolides, P. J., Perrone-Bizzozero, N. I., Finklestein, S. P. & Zwiers, H. *J. Neurosci.* **8**, 339-352 (1988).
- Oestreicher, A. B. & Gispén, W. H. *Brain Res.* **375**, 267-279 (1986).
- Griffin, J. W., Price, D. L., Drachman, D. B. & Morris, J. *J. Cell Biol.* **88**, 205-214 (1981).
- Pfenninger, K. H. & Maylie-Pfenninger, M.-F. *J. Cell Biol.* **89**, 547-559 (1981).
- Bambur, J. R. *Trends Neurosci.* **11**, 248-249 (1988).
- Patterson, P. H. *Neuron* **1**, 263-267 (1988).
- Lockerbie, R. O. *Neuroscience* **20**, 719-729 (1987).
- Alexander, K. S., Cimier, B. M., Meier, K. E. & Storm, D. R. *J. biol. Chem.* **262**, 6108-6113 (1987).
- van Dongen, C. J., Zwiers, H., de Graan, P. N. E. & Gispén, W. H. *Biochem. biophys. Res. Commun.* **128**, 1219-1227 (1985).
- van Hoof, C. O. M., de Graan, P. N. E., Oestreicher, A. B. & Gispén, W. H. *J. Neurosci.* **8**, 1789-1795 (1988).
- Kater, S. B., Mattson, M., Cohan, C. & Connor, J. *Trends Neurosci.* **11**, 315-321 (1988).
- Goldberg, D. J. *J. Neurosci.* **8**, 2596-2605 (1988).
- Lorentz, T. & Willard, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 505-509 (1978).
- Grafstein, B. & Forman, D. S. *Physiol. Rev.* **60**, 1167-1283 (1980).
- Banker, G. & Cowan, W. M. *Brain Res.* **126**, 397-425 (1977).
- Bartlett, W. P. & Banker, G. A. *J. Neurosci.* **4**, 1944-1953 (1984).

Selective localization of messenger RNA for cytoskeletal protein MAP2 in dendrites

Craig C. Garner*, Richard P. Tucker & Andrew Matus†

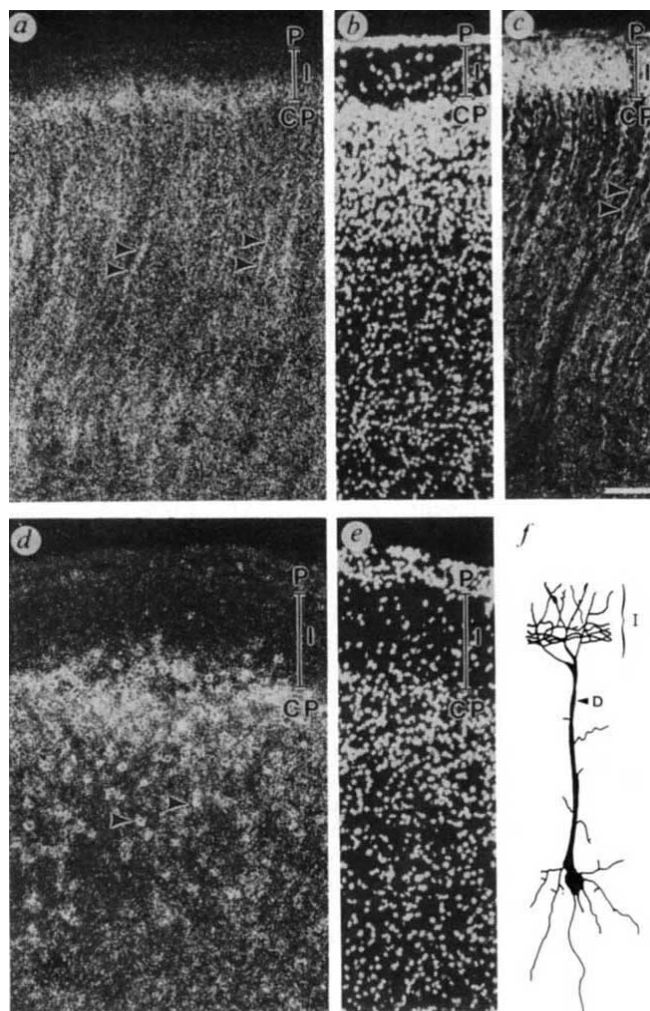
Friedrich Miescher-Institut, PO Box 2543, 4002 Basel, Switzerland

For nerve cells to develop their highly polarized form, appropriate structural molecules must be targeted to either axons or dendrites. This could be achieved by the synthesis of structural proteins in the cell body and their sorting to either axons or dendrites by specific transport mechanisms. For dendrites, an alternative possibility is that proteins could be synthesized locally in the dendritic cytoplasm. This is an attractive idea because it would allow regulation of the production of structural molecules in response to local demand during dendritic development. The feasibility of dendritic protein synthesis is suggested both by the existence of dendritic polyribosomes¹ and by the recent demonstration that newly synthesized RNA is transported into the dendrites of neurons differentiating in culture². However, to date there has been no demonstration of the selective synthesis of an identified dendrite-specific protein in the dendritic cytoplasm. Here, we use *in situ* hybridization with

* Present address: Zentrum für Molekulare Neurobiologie, 2000 Hamburg 20 FRG
† To whom correspondence should be addressed.

Fig. 1 The location of mRNAs for MAP2 (*a*) and tubulin (*d*) in six-day-old rat cerebral cortex revealed by *in situ* hybridization with cDNA probes. The location of label is shown by the white autoradiographic grains. The positions of cells in the same sections are shown by bisbenzidine fluorescent staining of nuclei (*b* and *e*). *c*, The distribution of MAP2 protein revealed by immunofluorescence staining. The pia (P) marks the outer edge of the cortex below which is the developing layer I (bracket marked "I"). Bundles of dendrites in *a* and *c* are indicated by arrowheads. Groups of tubulin-labelled cell bodies in *d* are indicated by single arrowheads; *a* and *c* show the same field from adjacent sections. *f*, A single developing pyramidal cell, traced from a drawing of Ramon y Cajal¹⁰, placed in relationship to the section in *e* to show the orientation of the apical dendrite (D) and terminal branching in layer I. Scale bar, 100 μ m.

Methods. Six-day-old rats were fixed by perfusion with 2% paraformaldehyde in potassium phosphate buffer, pH 7.2. The brains were removed and postfixed in the same solution at 4 °C overnight. Serial cryostat sections (12 μ m) were collected on gelatin-subbed slides, dried for 20 s on an 80 °C hotplate, air-dried for a further 1 h at room temperature and stored at -20 °C until required. Sections were thawed for 30 min at room temperature, pre-hybridized at room temperature in 4 $\times$ SSC, 1 $\times$ Denhardt's solution and 20 mM β -mercaptoethanol for 1 h and rinsed in ethanol. Hybridization with 0.1–1 ng of labelled cDNA probe was done overnight at 42 °C in 25 μ l of hybridization buffer (50% deionized formamide, 4 $\times$ SSC, 0.12M sodium phosphate, pH 7.2, 0.7% sarcosyl, 1 $\times$ Denhardt's solution and 1% dextran sulphate). Purified cDNA fragments encoding MAP2 (clone 19a, ref. 8) or β -tubulin⁹ were oligo-labelled with [α -³⁵S]dATP (400 Ci mol⁻¹; Amersham) as described²². After hybridization, sections were washed with 1 $\times$ SSC for 1 h at room temperature and for 1 h at 42 °C then dehydrated by two rinses in 0.33M ammonium acetate in 80% ethanol and two more rinses in 0.33M ammonium acetate in absolute ethanol. After air drying for 30 min at room temperature the slides were dipped in K5 nuclear emulsion (Ilford) diluted 1:1 in water, air-dried, and exposed at 4 °C for 1 month. After development with D-19 (Kodak) and fixing with Unifix (Kodak) slides were washed with water and stained with Hoechst nuclear dye (1 μ g ml⁻¹ bisbenzimidazole H 33258; Riedel-de-Haën) and photographed in dark field plus fluorescence on a Zeiss Universal microscope. Sections were immunofluorescence stained with monoclonal antibody AP14 against MAP2⁷.



specific complementary DNA probes to show that messenger RNA for the dendrite-specific microtubule-associated protein MAP2 (refs 3–5) is present in dendrites in the developing brain. By contrast the mRNA for tubulin, a protein present in both axons and dendrites^{4,6,7} is located exclusively in neuronal cell bodies.

Labelled cDNA probes for MAP2 (ref. 8) and tubulin⁹ gave different *in situ* hybridization patterns on sections of six-day-old rat brain. The MAP2 cDNA probe labelled the developing cerebral cortex with a pattern of vertically oriented stripes and a distinct band of labelling that was strongest in the inner half of layer I (Figs 1a and 2a). The vertical stripes (Fig. 1a) corresponded to bundles of MAP2-containing apical dendrites of developing pyramidal cells that could be seen by immunofluorescence staining with a monoclonal antibody against MAP2 (Fig. 1c). By contrast, the tubulin cDNA probe produced a distinct pattern of densely-labelled patches corresponding to clusters of cell bodies, the most strongly labelled cell groups being those that had most recently entered terminal differentiation at the outer edge of the cortical plate beneath layer I (Fig. 1d). The intense band of MAP2 cDNA labelling in layer I (Fig. 2a) corresponds to the area where the terminal branches of developing pyramidal cells are densely concentrated (Fig. 2c). There is no such band of labelling with the tubulin cDNA probe, which instead shows a sharp cut-off at the junction of the cortical plate with the layer I neuropil (compare Fig. 2b and d).

Similar differences between MAP2 and tubulin cDNA labelling patterns were also evident in other brain areas. In both the hippocampus and dentate gyrus, MAP2 labelling was intense throughout the molecular layers (Fig. 3a) where the dendrites

of pyramidal and granule cells are concentrated (Fig. 3d). The dendrites in these areas are not bundled but branch repeatedly producing a broad, dense band of cDNA labelling that corresponds to that seen for MAP2 protein by immunofluorescence (compare Fig. 3a and c). Tubulin cDNA labelling was associated with neuronal cell bodies, with both the pyramidal cell layer of the hippocampus and the granule cell layer of the dentate gyrus being strongly labelled (Fig. 2b). Instead of the strong, even labelling of the molecular layers, the tubulin pattern was marked by patches of grains corresponding to the cell bodies of interneurons.

These distinctive patterns suggest that tubulin mRNA is located in cell bodies whereas MAP2 mRNA is present in dendrites. The perikaryal location of tubulin mRNA is clear from its patchy distribution in the cortex, corresponding to clusters of cell bodies, its concentration in the pyramidal and granular cell layers of the hippocampus and its absence from neuropil areas between cells in both cortex and hippocampus. The dendritic localization of MAP2 mRNA is suggested by its distribution throughout the dendrite-rich neuropil areas of the cortex and hippocampus, and the similarity of the cDNA labelling patterns to that of MAP2 protein shown by immunofluorescence. Two distinctive features of the cortical MAP2 mRNA distribution are particularly suggestive of a dendritic localization. Firstly, the strong band of labelling in layer I corresponds to the location of the dendrite-rich neuropil areas of the developing apical dendrite at the external edge of the developing cortex and occupy the inner half of the developing layer I¹¹. By

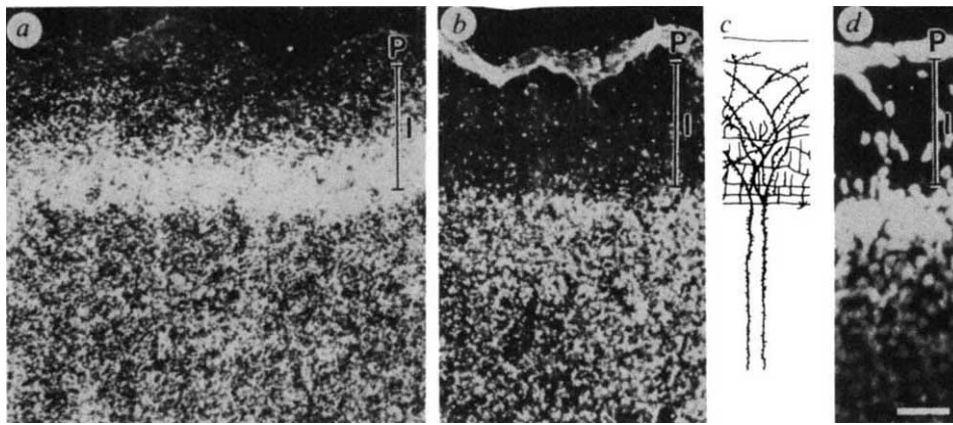


Fig. 2 MAP2 and tubulin cDNA hybridization patterns in the outer layer of the developing cortex. *a*, MAP2 probe; *b*, tubulin probe; *c*, location of apical dendrites as revealed by Golgi staining (tracing from example in ref. 11); *d*, location of cell bodies shown by fluorescent bisbenzimidazole nuclear dye. *P*, pial surface of cortex; bracket marked 'I' indicates layer I. Scale bar, 500 μ m.

contrast, there is little tubulin labelling in layer I. Second, the stripes that are labelled with the MAP2 probe correspond in size, orientation and location (between cell body columns) to bundles of pyramidal cell apical dendrites^{12,13}, which are a prominent feature of cortex sections stained with antibodies against MAP2 (Fig. 1c; see also ref. 14).

These results indicate that the dendritic localization of MAP2 that occurs throughout development^{6,7} is achieved by the selective association of MAP2 mRNA with the dendritic cytoplasm. This might indicate that MAP2 mRNA is transported to the developing dendrite from the pyramidal cell body. Alternatively MAP2 mRNA might be deposited in the apical dendrite, which is extruded from the cell body as it is displaced downwards from layer I by subsequently migrating neuroblasts¹. The second interpretation is supported by our observation of an intense band of MAP2 mRNA hybridization corresponding to the position of these apical dendrites. This would not exclude subsequent selective transport of MAP2 mRNA into dendrites as part of

normal mRNA turnover. Indeed, this is suggested by the transport of newly synthesized RNA into dendrites that has been observed in hippocampal pyramidal cells in culture² and the selective location of MAP2 in the dendrites of such cultured cells^{15,16}.

The presence of mRNA for MAP2 in dendrites suggests that the dendritic cytoplasm has an active role in controlling the synthesis of this protein during development. MAP2 promotes tubulin polymerization^{17,18} and is thought to be important in the regulation of dendrite development and plasticity^{19,20,21}. The location of its mRNA in the dendritic cytoplasm could allow the immediate control of MAP2 synthesis in response to local demands and thus make a significant contribution to the regulation of dendritic microtubule function and hence dendritic morphogenesis.

We thank Dr M. Kirschner for donating the tubulin cDNA clone used in this study and Dr L. I. Binder for the anti-MAP2 monoclonal antibody AP14.

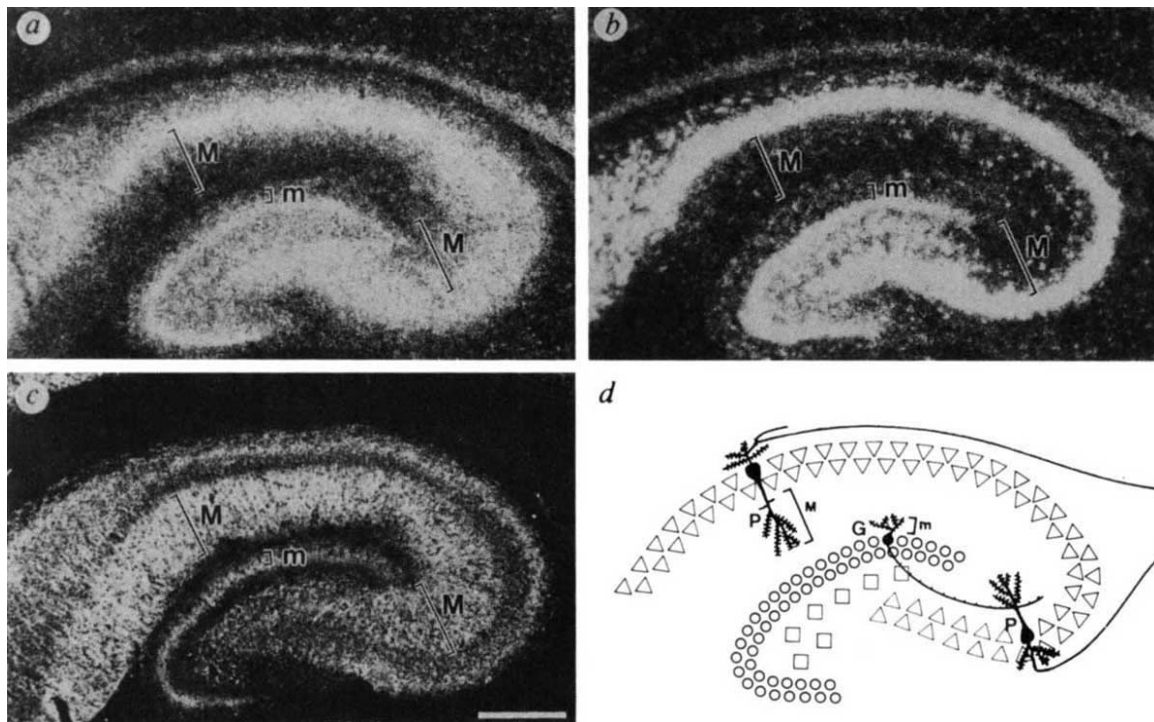


Fig. 3 *In situ* hybridization patterns of cDNA probes for MAP2 (*a*) and tubulin (*b*) in the hippocampus. *c*, The distribution of MAP2 protein revealed by immunofluorescent staining. Brackets marked 'M' indicate the stratum moleculare of the hippocampus, those marked 'm' the molecular layer of the dentate gyrus. *d*, The position of the pyramidal cell layer (triangles) and granule cells of the dentate gyrus (circles). The dendritic arrays of pyramidal cells (*P*) and granule cells (*G*) are indicated. Squares indicate the position of hilar neurons that are labelled by tubulin cDNA. Calibration bar, 250 μ m.

Received 23 September; accepted 21 October 1988.

1. Steward, O. & Falk, P. M. *J. Neurosci.* **6**, 412–423 (1986).
2. Davis, L., Banker, G. A. & Steward, O. *Nature* **330**, 477–479 (1987).
3. Bernhardt, R. & Matus, A. *J. comp. Neurol.* **226**, 203–221 (1984).
4. Caceres, A. *et al. J. Neurosci.* **4**, 394–410 (1984).
5. De Camilli, P., Miller, P. E., Navone, F., Theurkauf, W. E. & Vallee, R. B. *Neuroscience* **11**, 819–846 (1984).
6. Bernhardt, R. & Matus, A. *J. Cell Biol.* **92**, 589–593 (1982).
7. Tucker, R. P., Binder, L. I. & Matus, A. I. *J. comp. Neurol.* **271**, 44–55 (1988).
8. Garner, C. C. & Matus, A. *J. Cell Biol.* **106**, 779–783 (1988).
9. Cleveland, D. W. *et al. Cell* **20**, 95–105 (1980).
10. Cajal, S. R. *Histologie du Système Nerveux de l'Homme et des Vertébrés* Vol. 2, 519–646 (Maloine, Paris, 1952).
11. Marin-Padilla, M. & Marin-Padilla, T. M. *Anat. Embryol.* **164**, 161–206 (1982).
12. Fleischauer, K., Petsche, H. & Wittowski, W. Z. *Anat. EntwGesch.* **136**, 213–223 (1972).
13. Roney, K. J., Scheibel, A. B. & Shaw, G. *Brain Res. Rev.* **1**, 225–271 (1979).
14. Escobar, M. I. *et al. Neuroscience* **17**, 975–989 (1986).
15. Caceres, A., Banker, G., Steward, O., Binder, L. & Payne, M. *Dev. Brain Res.* **13**, 314–318 (1984).
16. Matus, A., Bernhardt, R., Bodmer, R. & Alaimo, D. *Neuroscience* **17**, 371–389 (1986).
17. Murphy, D. B. & Borisy, G. G. *Proc. natn. Acad. Sci. U.S.A.* **16**, (1977).
18. Herzog, W. & Weber, K. *Eur. J. Biochem.* **92**, 1–8 (1978).
19. Burgoyne, R. D. *Comp. Biochem. Physiol.* **83B**, 1–8 (1986).
20. Matus, A. A. *Rev. Neurosci.* **11**, 29–44 (1988).
21. Nunez, J. *Dev. Neurosci.* **8**, 125–144 (1986).
22. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).

Functional modulation of the nicotinic acetylcholine receptor by tyrosine phosphorylation

Jessica F. Hopfield*, David W. Tank†, Paul Greengard* & Richard L. Huganir*‡

* Laboratory of Molecular and Cellular Neuroscience, The Rockefeller University, New York, New York 10021, USA

† Molecular Biophysics Research Department, AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

Tyrosine-specific protein phosphorylation has been implicated in the regulation of cell transformation and proliferation^{1,2}. However, recent studies have shown that the expression of protein tyrosine kinases in adult brain is very high, suggesting that tyrosine-specific protein phosphorylation may also have a role in the regulation of neuronal function^{3–6}. Although a number of substrate proteins are phosphorylated on tyrosine residues, the functional alteration of proteins by tyrosine phosphorylation has previously been convincingly demonstrated only for protein tyrosine kinases^{7–9}. The nicotinic acetylcholine receptor, a neurotransmitter-gated ion channel, is phosphorylated by a protein tyrosine kinase in postsynaptic membranes *in vitro* and *in vivo*^{10,11}. We demonstrate here that this tyrosine phosphorylation increases the rate of the rapid phase of desensitization of the nicotinic receptor, as measured by single channel recording of purified nicotinic acetylcholine receptor, when reconstituted in lipid vesicles. These data provide direct evidence for the regulation of ion channel properties by tyrosine phosphorylation. The results, which demonstrate a functional role of tyrosine phosphorylation in the nervous system, suggest a widespread role for tyrosine phosphorylation in neuronal signal transduction.

The purified nicotinic acetylcholine receptor (nAChR) is a pentameric complex of four types of subunits arranged with stoichiometry $\alpha_2\beta\gamma\delta$ which can be functionally reconstituted into liposomes. To analyse quantitatively the properties of the purified and reconstituted nAChR from *Torpedo californica* and to determine the functional effects of tyrosine phosphorylation of the receptor, giga-seal patch-recording techniques were used to form isolated inside-out patches of reconstituted liposome membrane and to measure the single-channel properties of the receptor. The nAChR had channel properties similar to those

previously reported¹²: patches of liposome membrane containing the nAChR showed multiple channel openings corresponding to a unitary 4.2 pA current when the pipette contained 1 μ M acetylcholine and was held at +100 mV (Fig. 1a).

The frequency of acetylcholine-induced channel events diminished with time after seal formation and voltage application (Fig. 1a). This desensitization behaviour could be quantified from histograms of the number of channel openings per unit time after seal formation (Fig. 1b) and could be fitted to a sum of two exponentials. Two desensitization processes, a rapid and a slow phase, have previously been observed for the nicotinic acetylcholine receptor *in situ*^{13,14} and after reconstitution¹⁵. The rates of the rapid and slow phases of desensitization were dependent on the concentration of acetylcholine (see legend to Table 1) and were quantitatively similar to the rates of desensitization of the nAChR from *Torpedo californica* determined using rapid-mixing ion-flux techniques^{15,16}.

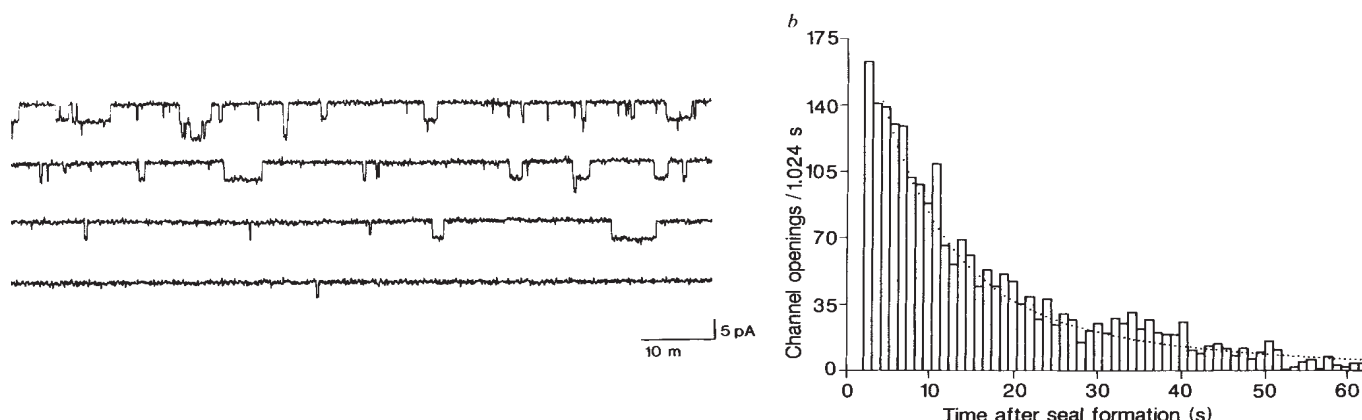
The nAChR from *T. californica* is phosphorylated on tyrosine residues to a high stoichiometry *in vivo* (unpublished results; Fig. 2). To determine the endogenous phosphotyrosine content of the purified nAChR we have applied quantitative immunoblotting techniques using an anti-phosphotyrosine affinity-purified antibody^{5,17} (Fig. 2). The purified nAChR contained 0.9–1.5 mol of endogenous phosphotyrosine per mol of receptor, depending on the preparation. The phosphotyrosine was distributed among the β , γ , and δ subunits (Fig. 2, lane 5). This endogenous phosphotyrosine content could be decreased *in vitro* by treatment of the nAChR with alkaline phosphatase (Fig. 2, lanes 3 and 4). Under the conditions used, alkaline phosphatase did not dephosphorylate nAChR that had been phosphorylated on serine residues *in vitro* by protein kinase C or the catalytic subunit of cAMP-dependent protein kinase, using [γ -³²P]ATP (data not shown). Conversely, the stoichiometry of tyrosine phosphorylation of the receptor could be increased *in vitro* by phosphorylation of the receptor using the endogenous protein tyrosine kinase activity in the postsynaptic membrane (Fig. 2, lanes 2, 6 and 7). Using a combination of these procedures, purified receptor preparations with eight different stoichiometries ranging from 0.6 to 2.7 moles phosphotyrosine per mole receptor were reconstituted into phospholipid vesicles, used for recording, and analysed for changes in channel kinetics.

The rate of the rapid phase of desensitization of the nAChR showed a striking dependence on the stoichiometry of tyrosine phosphorylation (Fig. 3, Table 1). Examples of data recorded from two individual patches of reconstituted receptor with different stoichiometries of tyrosine phosphorylation are shown in Fig. 3a and b. The rapid phase of desensitization of the receptor showed a sevenfold change in rate over the range of stoichiometries of tyrosine phosphorylation investigated (Fig. 3c). In addition, a twofold change in the slow phase of desensitization was observed (Table 1). Tyrosine phosphorylation of the nAChR had no detectable effect on other parameters of ion-channel function. The single-channel conductance and the apparent mean channel open-time distribution were essentially identical in all the receptor preparations and similar to those for untreated receptor¹².

Assuming that the phosphorylation/dephosphorylation of individual subunits occurs independently of one another, each nAChR preparation should consist of eight different receptor populations in varying proportions and showing a range in the number (0–3) of phosphates per receptor molecule and the individual subunits phosphorylated. The measured rates of desensitization presumably represent composites of the rates of desensitization of the various phosphorylated species. The data in Fig. 3 and Table 1 suggest that the effect of phosphorylation of each subunit is additive and that the more highly phosphorylated species of the nAChR desensitize more rapidly.

Most of the records from patches isolated from reconstituted nAChR preparations at the highest stoichiometries of tyrosine phosphorylation showed few channel events because of the

‡ To whom correspondence should be addressed. Present address: Howard Hughes Medical Institute and Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.



g. 1 Desensitization behaviour of purified and reconstituted nAChR as seen in single patches. *a*, 0.512 s portions of a single record sampled 15, 30 and 60 s following seal formation and voltage application. Pipette holding potential, +100 mV; 1 μ M acetylcholine. *b*, Exponential decay with time of the frequency of channel-opening events. The number of channel openings in successive 1.024 s intervals was plotted versus the midpoint of these time intervals. Histograms of records containing at least 100 channel events over the recording period were fitted to the following equation using a χ^2 minimization routine: $N(t) = [D(A/B e^{-t/B} + ((1-A)/C) e^{-t/C})] / [A e^{-t/B} + (1-A) e^{-t/C}]$. In this three-variable fit, *A* is % fast exponential's contribution to the observed exponential; *B* is τ_{fast} ; *C* is τ_{slow} ; *D* is total number of channel events $\times$ histogram bin width. The rapid phase of desensitization would itself be expected to be of multiexponential character, reflecting the multiple phosphorylation states of the receptor (see text). But the relatively low number of channel events per patch does not permit such complex exponential analysis. The denominator of the equation used is a correction for the distribution of channel events missed in the two-second delay before data collection. The dotted curve shows a double exponential distribution with $\tau_{fast} = 7.8$ s, $\tau_{slow} = 30.2$ s, % fast = 50.

Methods. The nAChR was purified and reconstituted as described³⁰, rapidly frozen and then thawed and diluted >3,000 times into a salt solution containing (mM) 150 NaCl, 1 CaCl₂, 1 MgCl₂, 10 HEPES, pH 7.2¹². The largest of the multilamellar liposomes (> 5.0 μ M in diameter) were selected under phase microscopy (400 $\times$) for recording. Patch pipettes (10–20 M Ω resistance) were filled with this salt solution plus varying concentrations of acetylcholine (300 nM–10 μ M). Two seconds after seal formation (seal resistance > 25 Gohm), the pipette holding potential was clamped at +100 mV and the pipette was withdrawn several microns from the liposome, creating a unilamellar inside-out patch. This two-second delay was required for seal stabilization. Currents were continuously sampled for 2 min at 4,000 Hz and filtered at 1,500 Hz (–3 dB, 8 pole Bessel filter). Experiments were conducted at room temperature (20–22 $^{\circ}$ C). The reconstituted nAChR displayed the established pharmacological responses to a variety of agonists and antagonists, including curare, procaine and α -bungarotoxin (data not shown). Desensitization of the nAChR was very rapid and channel events could be observed only if seal formation occurred within the first few seconds of the pipette entering the bath solution and if recording began immediately after seal formation. Such traces typically contained several hundred channel events during the two-minute period of continuous recording. Recordings obtained after longer delays to seal formation contained only a very low steady-state level of channel activity. Reconstituted receptors had a uniform orientation with the acetylcholine binding sites facing outward³⁰ and only patches with an inside-out configuration could be isolated in these experiments. It was therefore impossible to measure receptor recovery from desensitization because of the inability to rapidly remove reapplying agonist by internal perfusion of the patch-pipette. Open time analysis of nAChR in a single patch from records without simultaneous channel openings indicated that most channel lifetimes were in a range from <1–15 ms. The open channel distribution could be fitted to the sum of two exponentials with characteristic time constants of 0.9 ms and 7.0 ms with 83% of channels having the fast open time (data not shown).

Table 1 Effect of tyrosine phosphorylation of the nAChR on the time constants of desensitization

Total stoichiometry	Subunit stoichiometry			τ_{fast} (s)	τ_{slow} (s)	Per cent fast	<i>n</i>
	β	γ	δ				
0.6	0.2	0.1	0.3	15.3 (0.7)	36.3 (8.7)	87 (11)	4
0.9	0.3	0.3	0.3	9.7 (0.9)	23.1 (4.3)	86 (13)	3
1.5	0.6	0.3	0.6	7.2 (1.2)	23.6 (3.1)	45 (22)	6
1.5	0.5	0.5	0.5	6.9 (0.9)	25.5 (8.0)	71 (15)	6
1.8	0.6	0.6	0.6	5.4 (1.1)	20.6 (8.1)	87 (8)	7
2.0	0.7	0.5	0.8	2.9 (0.2)	19.9 (6.7)	44 (25)	6
2.3	0.7	0.7	0.9	2.7 (1.4)	19.1 (2.7)	40 (9)	4
2.7	0.9	0.9	0.9	2.2 (0.6)	14.8 (2.7)	55 (21)	3

Pipette holding potential +100 mV; 1 μ M acetylcholine. All experiments at a given stoichiometry (moles phosphotyrosine) were conducted with single preparation of reconstituted receptor. The rates of desensitization, measured at a stoichiometry of 0.5 mol phosphotyrosine per mole β , γ and δ subunits, showed the following dependence on the concentration of acetylcholine: 0.5 μ M: $\tau_{fast} = 11.5$ (0.8), $\tau_{slow} = 28.8$ (3.9), % fast = 61 (25), 1.0 μ M: see table; 2.5 μ M: $\tau_{fast} = 3.9$ (0.8), $\tau_{slow} = 20.9$ (1.5), % fast = 68 (19), *n* = 4; 5.0 μ M: $\tau_{fast} = 1.6$ (0.5), $\tau_{slow} = 12.0$ (2.2), % fast = 79 (22). All data are mean values ($\pm$ standard deviations).

very rapid rate of desensitization. In such cases, most of channel openings appear to have occurred during the several-second delay before data collection, which is inherent in the recording paradigm. But when patches were isolated with an unusually short delay, a significant portion of the rapid phase of desensitization could be observed. For the most highly phos-

phorylated preparation used (stoichiometry, 2.7 mol phosphotyrosine per mol receptor), 73% of the receptor molecules would be expected to be phosphorylated on the β , γ and δ subunits. It is possible that the two-second time constant observed may be largely determined by the 27% of the nAChR molecules not phosphorylated on all three subunits and,

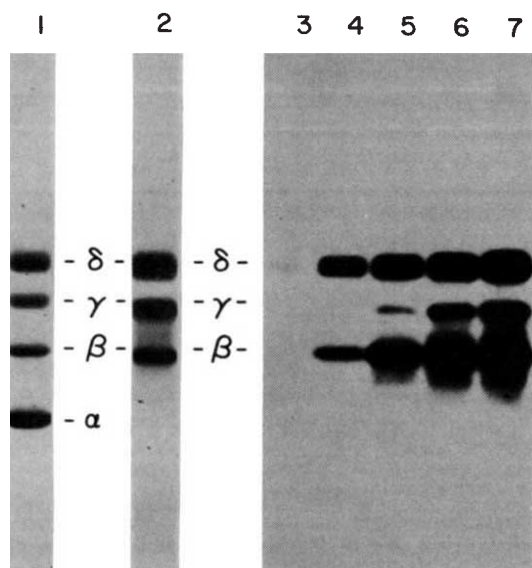


Fig. 2 Purified nAChR preparations with various stoichiometries of tyrosine phosphorylation: lane 1, protein stain of purified receptor preparation; lane 2, autoradiograph of ^{32}P -labelled tyrosine phosphorylated receptor; lanes 3–7, Western blot analysis of nAChR preparations with increasing phosphotyrosine content using affinity-purified anti-phosphotyrosine antibody⁵, lanes 3 and 4, *in vitro* phosphatase-treated nAChR; lane 5 untreated receptor; lanes 6 and 7, *in vitro* phosphorylated nAChR.

Methods. Postsynaptic membranes highly enriched in the nAChR were isolated from the electric organ of *T. californica*¹⁰. The postsynaptic membranes were diluted to 0.5 mg ml^{-1} with phosphorylation reaction mixture¹⁰ with (lanes 6 and 7) or without (lanes 3–5) 0.5 mM ATP . Lane 2 was incubated with $0.5 \text{ mM } [\gamma\text{-}^{32}\text{P}]\text{ATP}$. Under these conditions, the nAChR is phosphorylated exclusively on tyrosine residues. The incubation was carried out at 30°C for $30'$ (lanes 2–5, 7) or for $10'$ (lane 6). The reactions were stopped by rapid cooling to 4°C and centrifugation of the phosphorylated membranes. The pellets were resuspended at 2.5 mg ml^{-1} in 20 mM Tris-HCl , $\text{pH } 8.0$, 20 mM MgCl_2 , 1 mM EDTA , 1 mM EGTA , $20 \mu\text{g ml}^{-1}$ leupeptin, $20 \mu\text{g ml}^{-1}$ antipain, 20 units ml^{-1} Trasylol, 0.3% sodium cholate. The suspensions were incubated for $15'$ at 30°C with alkaline phosphatase (Boehringer Mannheim) at 0.25 mg ml^{-1} (lane 4) and 0.5 mg ml^{-1} (lane 3), or without alkaline phosphatase (lanes 2, 5–7). The reaction mixture was cooled to 4°C . The nAChR was then purified on acetylcholine affinity columns and reconstituted as previously described³⁰. To determine the stoichiometry of phosphotyrosine content of the nAChR subunits, Western blots were performed using affinity-purified anti-phosphotyrosine antibodies and ^{125}I -labelled protein A. To convert ^{125}I -protein A labelling of the nAChR to mol of phosphotyrosine, the nAChR was phosphorylated *in vitro* using either unlabelled ATP or $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ of known specific activity. The specific activity of ^{125}I -protein A labelling (c.p.m. per mol phosphotyrosine) was calculated by comparing the stoichiometry of ^{32}P incorporated with the increase in ^{125}I -protein A labelling of the nAChR after phosphorylation *in vitro*, using the formula [(c.p.m. of ^{125}I -protein A labelling of subunit phosphorylated *in vitro*) – (c.p.m. of ^{125}I -protein A labelling of nonphosphorylated subunit) per mol ^{32}P -phosphate incorporated *in vitro*]. From this calculation, the stoichiometry of tyrosine phosphorylation of the subunits of all the receptor preparations was determined.

therefore, the rate constant of the rapid phase of desensitization for the triply phosphorylated receptor may be considerably faster than the two seconds measured for this preparation.

The results presented here demonstrate that tyrosine phosphorylation of the nAChR dramatically increases the rate of the rapid phase of desensitization. Moreover, dephosphorylation of the endogenous phosphotyrosine residues decreases the rate of the rapid phase of desensitization to the point where it

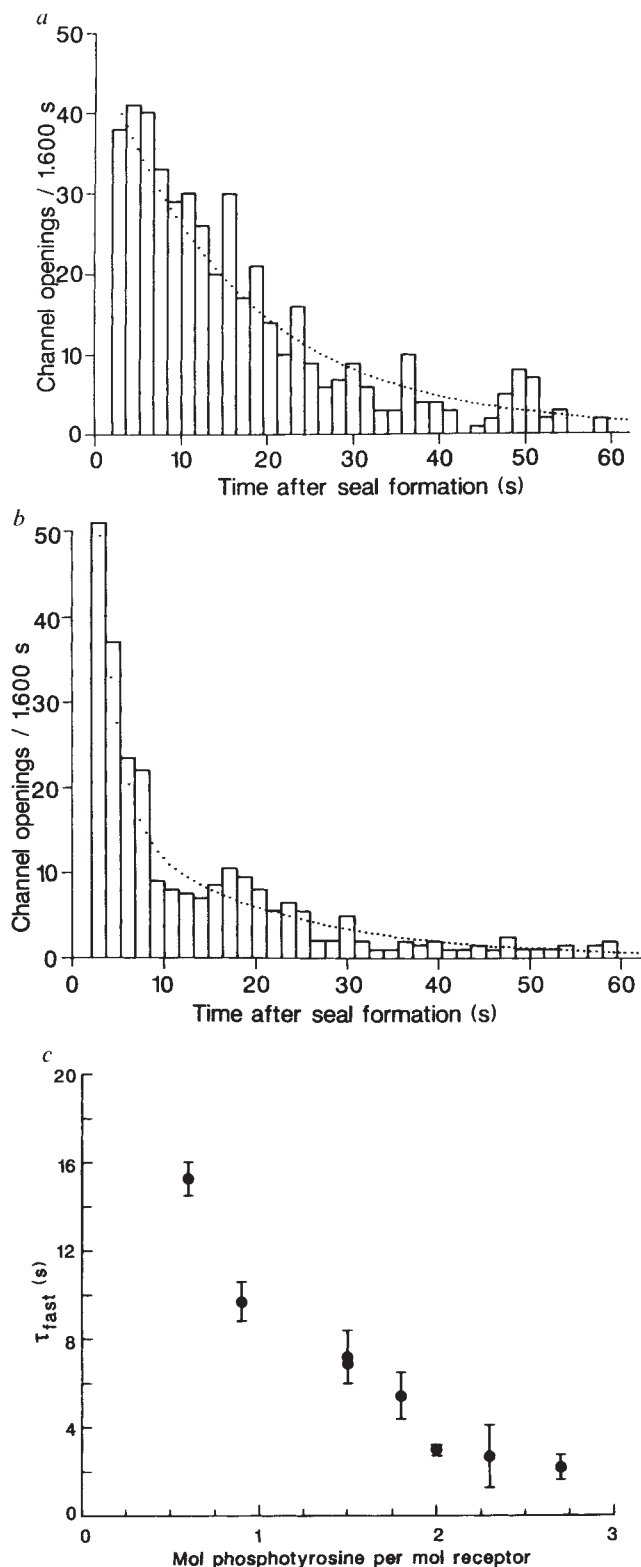


Fig. 3 Increased levels of tyrosine phosphorylation increase the rate of the rapid phase of desensitization of the nAChR. Pipette holding potential $+100 \text{ mV}$; $1 \mu\text{M}$ acetylcholine. **a**, Exponential decay with time in the frequency of channel opening events, $0.6 \text{ mol phosphotyrosine per mol receptor}$. Fitted double exponential: $\tau_{\text{fast}} = 15.1 \text{ s}$, $\tau_{\text{slow}} = 32.4 \text{ s}$, % fast = 76 . **b**, Exponential decay with time in the frequency of channel opening events, $2.3 \text{ mol phosphotyrosine per mol receptor}$. Fitted double exponential: $\tau_{\text{fast}} = 2.1 \text{ s}$, $\tau_{\text{slow}} = 17.6 \text{ s}$, % fast = 47 . **c**, Time constants of the rapid phase of desensitization are plotted versus the stoichiometry of tyrosine phosphorylation. Data are plotted as mean values $\pm$ standard deviations.

approaches the rate of the slow phase of desensitization. These results indicate that the rapid phase of desensitization could have an absolute dependence on phosphorylation. We have been unable to test this hypothesis because we have not been able to prepare a completely dephosphorylated preparation of the nAChR.

The nAChR is phosphorylated by three different protein kinases: cyclic AMP-dependent protein kinase¹⁸, protein kinase C^{19,20} and a tyrosine-specific-protein kinase¹⁰. Cyclic AMP-dependent phosphorylation of the γ and δ subunits of the purified nAChR increases the rate of the rapid phase of desensitization¹⁶. In addition, studies with intact muscle^{21,22} and cultured muscle cells^{23,24} have shown that cAMP analogues and forskolin (both activators of cAMP-dependent protein kinase) and phorbol esters (activators of protein kinase C) increase the rate of desensitization of the receptor. These results, taken together with those presented here, suggest that phosphorylation of the receptor by each of the three protein kinase systems, presumably in response to three different first messenger pathways, regulates the rate of desensitization of the nAChR. The first messengers that regulate phosphorylation of the nAChR by protein kinase C and the tyrosine-specific protein kinase are not known. However it has been reported recently that calcitonin gene-related peptide (CGRP), a neuropeptide that is present at the neuromuscular junction, regulates phosphorylation of the nAChR²⁵ and the rate of desensitization of the nAChR in myotubes²⁶ by the activation of cAMP-dependent protein kinase. Desensitization is a form of short-term regulation of synaptic efficacy in the second-to-minute time range²⁷⁻²⁹. Protein phosphorylation appears to be an important way of modulating this process.

We thank Lisa Moritz for technical assistance. Jessica F. Hopfield holds an NSF predoctoral fellowship. This work was

supported by USPHS grants to Richard L. Haganir and Paul Greengard.

Received 14 September; accepted 20 October 1988.

- Hunter, T. & Cooper, J. A. *Rev. Biochem.* **54**, 897-930 (1983).
- Sefton, B. M. *Curr. Topics microbiol. Immun.* **123**, 40-72 (1986).
- Cotton, P. C. & Brugge, J. S. *Molec. cell. Biol.* **3**, 1157-1162 (1983).
- Sudol, M. & Hanafusa, H. *Molec. cell. Biol.* **6**, 2839-2846 (1986).
- Hirano, A. A., Greengard, P. & Haganir, R. L. *J. Neurochem.* **50**, 1447-1455 (1988).
- Pang, D. T., Wang, J. K. T., Valtorta, F., Benfenati, F. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 762-766 (1988).
- Rosen, O. M., Herrera, R., Olowe, Y., Petruzzelli, J. M. & Cobb, M. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3237-3240 (1983).
- Courtneidge, S. A. *EMBO J.* **4**, 1471-1477 (1985).
- Cooper, J. A. & King, C. S. *Molec. cell. Biol.* **6**, 4467-4477 (1986).
- Haganir, R. L., Miles, K. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6968-6972 (1984).
- Smith, M. M., Merlie, J. P. & Lawrence, J. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6601-6605 (1987).
- Tank, D. W., Haganir, R. L., Greengard, P. & Webb, W. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5129-5133 (1983).
- Sakmann, B., Patlak, J. & Potter, L. T. *Nature* **286**, 71-73 (1980).
- Hamill, O. P. & Sakmann, B. *Nature* **286**, 462-464 (1981).
- Walker, J. W., Jakeyasu, K. & McNamee, M. G. *Biochemistry* **21**, 5384-5389 (1982).
- Haganir, R. L., Delcour, A. H., Greengard, P. & Hess, G. P. *Nature* **321**, 774-776 (1986).
- Ohtsuka, M., Ihara, S., Ogawa, R., Watanabe, T. & Watanabe, Y. *Int. J. Cancer* **34**, 855-861 (1984).
- Haganir, R. L. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1130-1134 (1983).
- Haganir, R. L., Albert, K. A. & Greengard, P. *Soc. Neurosci. Abstr.* **9**, 578 (1983).
- Safran, A., Sagi-Eisenberg, R., Neumann, D. & Fuchs, S. *J. biol. Chem.* **262**, 10506-10510 (1987).
- Middleton, P., Jaramillo, F. & Scheutze, S. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4967-4971 (1986).
- Albuquerque, E. X., Deshpande, S. S., Aracava, Y., Alkonon, M. & Daly, J. W. *FEBS Lett.* **199**, 113-120 (1986).
- Eusebi, F., Molinaro, M. & Zani, B. M. *J. Cell Biol.* **100**, 1339-1342 (1985).
- Middleton, P., Rubin, L. L. & Schuetze, S. M. *J. Neurosci.* **8**, 3405-3412 (1988).
- Miles, K., Greengard, P. & Haganir, R. L. *Neuron* (submitted).
- Mulle, C., Benoit, P., Pinset, C., Roa, M. & Changeux, J.-P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5728-5732 (1988).
- Krnjević, K. & Miledi, R. *J. Physiol.* **140**, 440-461 (1958).
- Magleby, K. S. & Pallotta, B. S. *J. Physiol.* **316**, 225-250 (1981).
- Changeux, J.-P., Devillers-Thiery, A. & Chemouilli, P. *Science* **225**, 1335-1345 (1984).
- Haganir, R. L. & Racker, E. *J. biol. Chem.* **257**, 9372-9378 (1982).

The G protein-gated atrial K⁺ channel is stimulated by three distinct G_iα-subunits

Atsuko Yatani*, Rafael Mattera‡, Juan Codina†, Rolf Graf†, Kouji Okabe*, Elena Padrell§, Ravi Iyengar§, Arthur M. Brown* & Lutz Birnbaumer*†

Departments of Physiology and Molecular Biophysics*, Cell Biology†, and Medicine‡, Baylor College of Medicine, Houston, Texas 77030, USA

§ Department of Pharmacology, Mount Sinai School of Medicine, New York, New York 10029, USA

The guanine nucleotide-binding protein, G_i, which inhibits adenylyl cyclase, has recently been shown to have three subtypes of the α-subunit, termed G_iα-1, G_iα-2 and G_iα-3. They share 87-94% amino-acid sequence homology¹⁻¹⁰ and so are difficult to separate from one another. Among other functions^{11,12-14}, purified preparations activate K⁺ channels¹⁵⁻²¹ but there is confusion over which of the subtypes activates the muscarinic K⁺ channels of the atrial muscle of the heart: G_iα-3, also termed G_k¹⁵, has been shown to activate this channel but it is not clear whether G_iα-1²² does²¹ or does not^{23,24}. To clarify this problem, we expressed the subtypes separately in *Escherichia coli* to eliminate contamination by other subtypes and tested the recombinant α-chains on atrial muscarinic K⁺ channels. Although we anticipated that only G_iα-3 would have G_k activity, to our surprise all three recombinant subtypes were active, from which we deduce that the G_i subtypes are multifunctional.

We modified the Gα complementary DNAs by site-directed mutagenesis, if necessary, to retain a unique C'CATGG *Nco*I site at the start of their open reading frames and inserted the

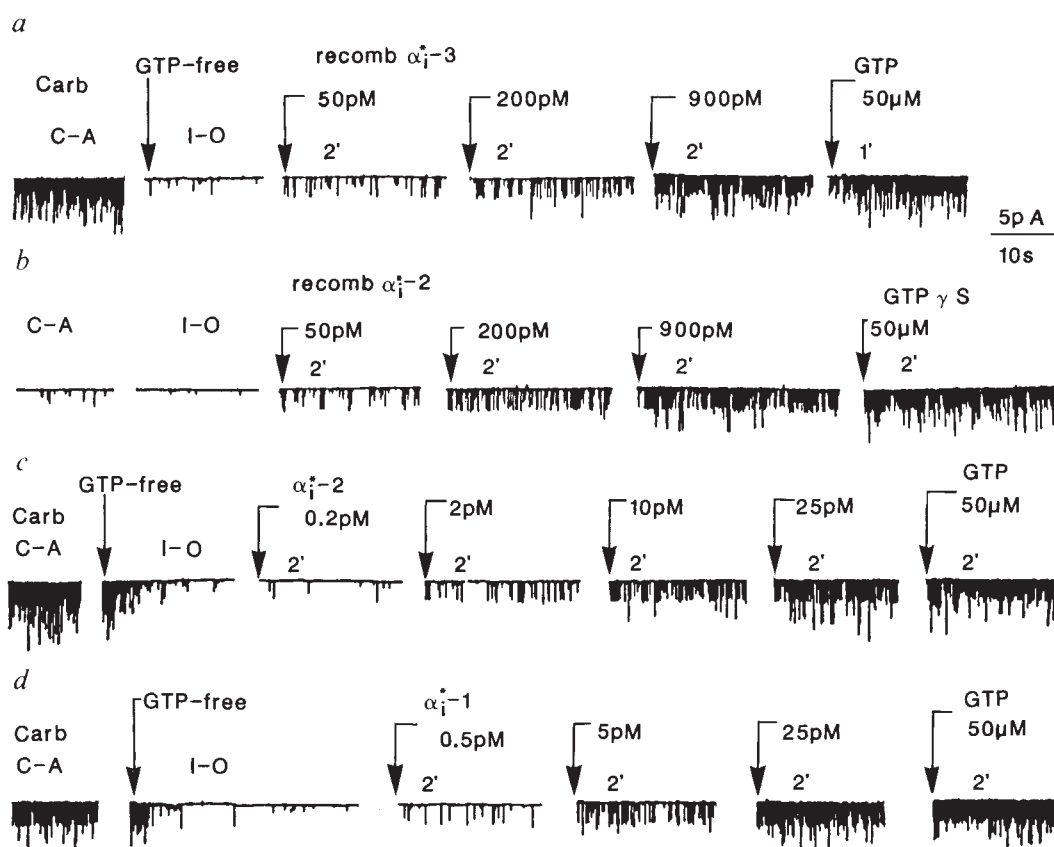
cDNAs into the *Bam*HI site of the cloning cassette of the pT7-7 expression vector²⁵. This construction led to the expression of recombinant Gα fusion polypeptides that differed from their natural counterparts by an N-terminal extension of nine amino acids. The recombinants constituted 5-10% of the total cell protein²⁶ of which 5% was recovered in a GTPγS-activated form after partial purification by DEAE-cellulose chromatography²⁶.

We found that the GTPγS-activated recombinant forms of G_iα-3, G_iα-1 and G_iα-2 stimulated specific atrial K⁺ channels (Figs 1 and 2). The conductances and mean open times were identical to those of atrial K⁺ channels stimulated by carbachol plus GTP or GTPγS (Table 1). Maximal activation by any one of the recombinant G_iα-subunits was not additive with stimulation by agonist plus GTP or GTPγS (Fig. 1). The stimulation was concentration dependent: concentration-response curves were constructed to compare intrinsic potencies. As for activated native human erythrocyte G_iα-3¹⁵, the stimulatory effects of the GTPγS-activated recombinant G_iα-subunits could not be washed out and were essentially irreversible (data not shown). Hence, the effectiveness of added G_iα-subunits depended on concentration and time. Currents produced at each concentration were compared from about 120 s after subunit addition for the next 20 s (Fig. 2). For normalization, the maximal patch currents were measured with 100 nM carbachol in the pipette and 50-100 μM GTP in the bath, or after addition of 50-100 μM GTPγS to the bath. Under these conditions we found the three forms of recombinant G_iα-subunits were equipotent (Fig. 2) with one difference: the frequency of success with recombinant G_iα-1 seemed lower. Our impression is that failures result when access to the patch is restricted due to partial sealing over. This seems to affect recombinant G_iα-1 more than recombinants G_iα-2 and G_iα-3.

The following tests proved that the effects observed with each of the recombinant G_iα-subunits were intrinsic properties of

Fig. 1 Effects of GTP γ S-activated $G_i\alpha$ -2 and $G_i\alpha$ -3 recombinant subunits (recomb α_i^*-3 (a) and recomb α_i^*-2 (b)), and of natural GTP γ S-activated $G_i\alpha$ -subunits (α_i^*-2 , c; α_i^*-1 , d) on single atrial muscarinic K^+ -channel currents of inside-out membrane patches excised from single atrial cells of adult guinea pigs¹⁵. The effects of recombinant α_i^*-1 are shown in Fig. 2.

Methods. Recombinant α -subunits were induced in *E. coli*, extracted, activated with GTP γ S and purified partially, as previously described²⁶. Natural G_i subunits were prepared¹⁶ from bovine brain $G_i\alpha$ -1 (E. Padrell & R. Iyengar, manuscript in preparation) and human erythrocyte $G_i\alpha$ -2²⁸. The proteins were diluted in bathing solution¹⁵ and added to inside-out membrane patches as described¹⁶. Single channel K^+ currents were recorded and analysed as described^{15,16,27}. In a, c



and d, carbachol (Carb) was in the pipette at 10^{-7} M. The effects are compared to those of GTP (in the presence of carbachol) or GTP γ S (in the absence of agonist, b) in each experiment. The results were unaffected by the presence or absence of carbachol in the pipette.

the expressed molecules. (1) Boiling before addition to the bath abolished the activity of each recombinant $G_i\alpha$ -subunit ($n = 6$). (2) Partially-purified proteins from bacteria transfected with the expression plasmid having antisense cDNAs or partially purified, GTP γ S-activated recombinant $G_i\alpha$ -subunit which stimulates adenylyl cyclase and dihydropyridine-sensitive Ca^{2+} channels²⁷, had no effect ($n = 18$). (3) Threshold concentrations of GTP γ S for activation of atrial muscarinic K^+ channels in inside-out membrane patches with 100 nM carbachol in the pipette were 10–100 nM, at least 10 times the maximum GTP γ S added with saturating concentrations (1 nM) of any of the GTP γ S-activated recombinant $G_i\alpha$ -subunits. Prior addition of

100 μ M GDP β S, which blocked carbachol-mediated effects of 10 μ M GTP, did not interfere with the effects of the GTP γ S-activated recombinant $G_i\alpha$ -subunits.

Prompted by the results with the $G_i\alpha$ -1 and $G_i\alpha$ -2 recombinant subunits, we purified natural $G_i\alpha$ -1 from bovine brain (E. Padrell and R. Iyengar, unpublished results) and $G_i\alpha$ -2 from human erythrocytes²⁸, prepared their GTP γ S-activated forms¹⁶ and confirmed that both had potent G_k activity (Fig. 1). Half-maximal effects were obtained with approximately 10–30 pM of $G_i\alpha$ -1 and $G_i\alpha$ -2 (not shown), values not significantly different from those previously obtained with human erythrocyte $G_i\alpha$ -3^{15,16,19}.

Figure 2 shows that the concentrations of recombinant $G_i\alpha$ -subunits needed for half-maximal effects were approximately 500 pM, about 25 times higher than the half-maximal concentrations for the native $G_i\alpha$ -subunits. The maximal currents produced, however, were similar. Another bacterial fusion polypeptide, a recombinant $G_s\alpha$ -subunit, stimulates adenylyl cyclase and dihydropyridine-sensitive Ca^{2+} channels with less potency than the native protein²⁸. Perhaps incomplete or aberrant post-translation processing is responsible for the lower affinity of the recombinant polypeptides for their effectors.

These experiments showed that native $G_i\alpha$ -1, $G_i\alpha$ -2 and $G_i\alpha$ -3 reproduced the molecular behaviour of the natural atrial G_k protein (Table 1). Now, as before²⁹, it is unclear why bovine brain GTP γ S-activated G_{41} repeatedly failed to activate the atrial muscarinic K^+ channel^{23,24}. Positive results would be expected regardless of the exact identity ($G_i\alpha$ -1, $G_i\alpha$ -2 or $G_i\alpha$ -3) of the pertussis toxin substrate used in those experiments.

The demonstration that all three $G_i\alpha$ -subunits are isoforms with respect to K^+ channel regulation leads us to conclude that one or more of these G proteins must be multifunctional. Previous work has shown that another G protein, G_s , is multifunctional, and can stimulate both adenylyl cyclase and dihy-

Table 1 Comparison of single atrial K^+ -channel currents activated by carbachol plus GTP, GTP γ S and the three types of GTP γ S-activated recombinant $G_i\alpha$ -subunits

Conditions	Mean open time (ms)	Conductance (pS)	n
Agonist + GTP	1.8 ± 0.6	40 ± 2	18
GTP γ S	1.8 ± 0.8	40 ± 1	8
Recombinant $G_i\alpha$ -1	1.6 ± 0.5	40 ± 1	7
Recombinant $G_i\alpha$ -2	1.7 ± 0.7	40 ± 2	6
Recombinant $G_i\alpha$ -3	1.8 ± 0.5	39 ± 2	8

Data were from inside-out membranes patches excised from adult guinea pig atrial cells. Solutions were symmetrical isotonic K^+ solutions¹⁵. When present, carbachol was at 10^{-7} M; GTP at 10–100 μ M; GTP γ S at 50–100 μ M; and recombinant $G_i\alpha$ -subunits, at 0.2–1 nM. In some experiments AMP-P(NH)P 200 μ M was present. Single-channel currents were analysed with a laboratory computer (PDP 11/73, Digital Equipment Corp.) as described previously^{15,16,27}. Slope conductances were measured between -40 and -120 mV in half the experiments and chord conductances were measured at -40 mV in the remainder. n is the number of experiments and values are expressed as means $\pm$ s.e.m.

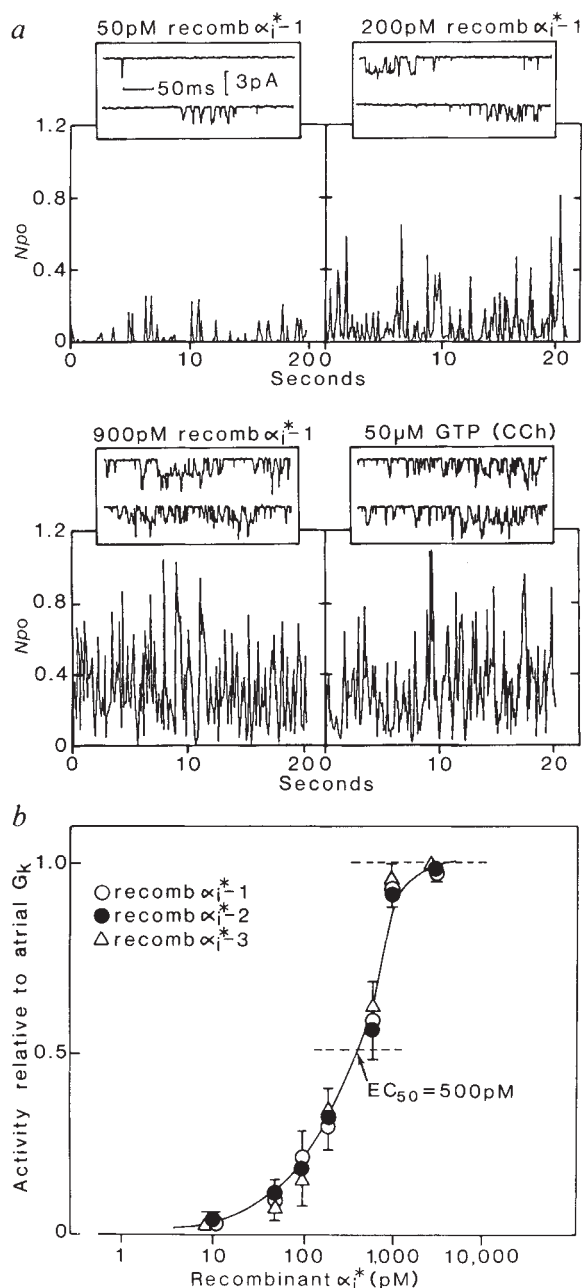


Fig. 2 Concentration-response relationships for recombinant α_i^* -subunits. In *a*, recombinant $G_i\alpha^*-1$ (recomb α_i^*-1) was added at 50, 200 and 900 pM. Insets are the single channel currents at -80 mV in symmetrical K^+ solutions^{15,16}. Filter frequency at -3 dB was 1 kHz and sampling rate was 5 kHz. The proportion of open time was measured every 100 ms and plotted as Np_0 (N is the number of channels; p_0 , the probability of the channel being open at that potential) versus time, to give the diaries shown in the main panels. Np_0 was averaged over 200 such segments beginning in each case 120 s after addition of recombinant α_i^*-1 and the mean values were 0.028, top left, 0.0903 top right and 0.319 (bottom left). The addition of 100 μ M GTP to the bathing solution elicited no response beyond that obtained with 900 pM recombinant $G_i\alpha^*-1$ (bottom right mean value was 0.320). In *b*, the mean values obtained with recombinant $G_i\alpha^*-1$, $G_i\alpha^*-2$ and $G_i\alpha^*-3$, were normalized to maximum Np_0 values obtained either by adding 50–100 μ M GTP when carbachol at 10^{-7} M was in the pipette or by adding 50–100 μ M GTP γ S when carbachol was absent, and were plotted against subunit concentrations. Data points were mean $\pm$ s.e.m. from 4 to 6 patches. The concentration-response relationships obtained for each of the three types of subunits were superimposable, had a Hill coefficient greater than one and reached 50% of maximum at 500 pM (EC_{50}) of added recombinant $G_i\alpha^*$ -subunits.

dropyridine-sensitive Ca^{2+} channels^{26,27,30}, indicating that many G proteins may be multifunctional. In the world of signal transduction, a G protein does not simply regulate one effector; it executes a programme of responses.

Received 17 August; accepted 3 October 1988.

1. Nukada, T. *et al.* *FEBS Lett.* **197**, 305–310 (1986).
2. Itoh, *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 3776–3780 (1986).
3. Suki, W., Abramowitz, J., Mattera, R., Codina, J. & Birnbaumer, L. *FEBS Lett.* **220**, 187–192 (1987).
4. Itoh, H., Toyama, R., Kozasa, T., Tsukamoto, T., Matsuoka, M. & Kaziro, Y. *J. biol. Chem.* **263**, 6656–6664 (1988).
5. Sullivan, K. A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 6687–6691 (1986).
6. Didsbury, J. R., Ho, Y.-S. & Snyderman, R. *FEBS Lett.* **211**, 160–164 (1987).
7. Didsbury, J. R. & Snyderman, R. *FEBS Lett.* **219**, 259–263 (1987).
8. Jones, D. T. & Reed, R. R. *J. biol. Chem.* **262**, 14241–14249 (1987).
9. Beals, C. R., Wilson, C. B. & Perlmutter, R. M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7886–7890 (1987).
10. Codina, J. *et al.* *J. biol. Chem.* **263**, 6746–6750 (1988).
11. Roof, D. J., Applebury, M. L. & Sternweis, P. C. *J. biol. Chem.* **260**, 16242–16249 (1985).
12. Kikuchi, A. *et al.* *J. biol. Chem.* **261**, 11558–11562 (1986).
13. Hescheler, J. *et al.* *EMBO J.* **7**, 619–624 (1988).
14. Rosenthal, W. *et al.* *EMBO J.* **7**, 1627–1633 (1988).
15. Yatani, A., Codina, J., Brown, A. M. & Birnbaumer, L. *Science* **235**, 207–211 (1987).
16. Codina, J., Yatani, A., Grenet, D., Brown, A. M. & Birnbaumer, L. *Science* **236**, 442–445 (1987).
17. Yatani, A., Codina, J., Sekura, R. D., Birnbaumer, L. & Brown, A. M. *Molec. Endocrinol.* **1**, 283–289 (1987).
18. Codina, J., Grenet, D., Yatani, A., Birnbaumer, L. & Brown, A. M. *FEBS Lett.* **216**, 104–106 (1987).
19. Kirsch, G., Yatani, A., Codina, J., Birnbaumer, L. & Brown, A. M. *Am. J. Physiol.* **254**, H1200–H1205 (1988).
20. Cerbai, E., Kloeckner, U. & Isenberg, G. *Science* **240**, 1782–1784 (1988).
21. Szabo, G., Pang, I. H. & Sternweis, P. C. *Biophys. J.* **53**, 424a (1988).
22. Memby, S., Pang, I.-H., Gilman, A. G. & Sternweis, P. C. *J. biol. Chem.* **263**, 2020–2026 (1988).
23. Logothetis, D. E., Kurachi, Y., Galper, J., Neer, E. J. & Clapham, D. E. *Nature* **325**, 321–326 (1987).
24. Logothetis, D. E., Kim, D., Northup, J. K., Neer, E. J. & Clapham, D. E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5814–5818 (1988).
25. Tabor, S. thesis, Harvard Univ. (1987).
26. Mattera, R. *et al.* *Science* (submitted).
27. Yatani, A. *et al.* *J. biol. Chem.* **263**, 9882–9895 (1988).
28. Birnbaumer, L. *et al.* *Cold Spring Harb. Symp. quant. Biol.* (in the press).
29. Birnbaumer, L. & Brown, A. M. *Nature* **327**, 21–22 (1987).
30. Yatani, A. *et al.* *Science* **238**, 1288–1292 (1987).

Conservation of the D-mannose-adhesion protein among type 1 fimbriated members of the family Enterobacteriaceae

Soman N. Abraham, Daxi Sun, James B. Dale & Edwin H. Beachey*

The Veterans Administration Medical Center and the University of Tennessee, Memphis, Tennessee 38104, USA

A variety of genera and species of the family Enterobacteriaceae bear surface fimbriae that enable them to bind to D-mannose residues on eukaryotic cells^{1–3}. Until recently, it was thought that the D-mannose binding site was located in the major structural subunit (FimA), of relative molecular mass (M_r) 17,000 (17 K), of these organelles in *Escherichia coli*^{4,5}. New evidence indicates that this binding site resides instead in a minor protein M_r 28–31 K (FimH) located at the tips and at long intervals along the length of the fimbriae^{6–12}, and is reminiscent of the minor tip adhesion proteins of pyelonephritis-associated pili (Pap) and S fimbriae^{13,14}. In contrast to the antigenic heterogeneity of the major FimA subunit, the antigenic structure of FimH is conserved among different strains of *E. coli*^{10,11}. Here, we report an even broader conservation of this minor adhesion protein extending to other genera and species of type 1 fimbriated Enterobacteriaceae. Our results may have implications for the development of broadly protective vaccines against Gram-negative bacillary infections in animals and perhaps in man.

Type 1 fimbriae were isolated from two random clinical isolates each of *E. coli*, *Klebsiella pneumoniae*, *Serratia marcescens*

* To whom correspondence should be addressed.

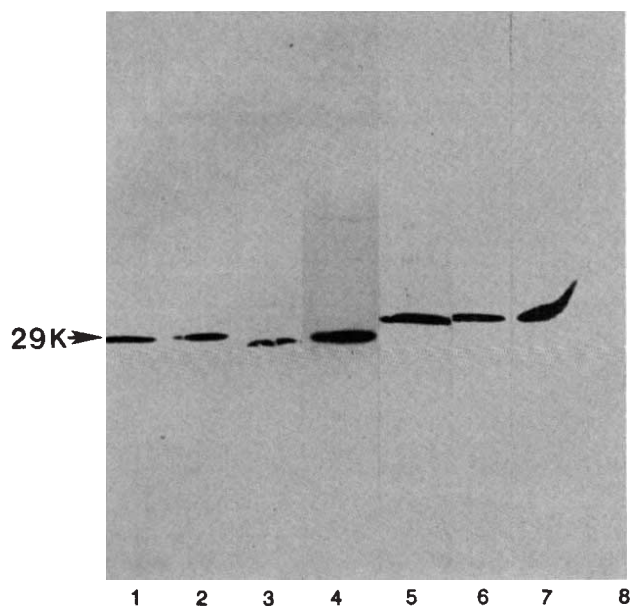


Fig. 1 Western blot analysis of isolated type 1 fimbriae from four species of the family Enterobacteriaceae. Lanes 1-7 represent dissociated fimbriae from various bacterial strains that were electrophoresed on a 15% SDS gel, transferred onto nitrocellulose paper and reacted with monoclonal antibody IIF4: lane 1, *E. coli* CI5; 2, *E. coli* CI 1400; 3, *K. pneumoniae* 1,376; 4, *K. pneumoniae* 59; 5, *S. marcescens* 1,495; 6, *S. marcescens* 50; 7, *C. freundii* 39; 8, control Pap fimbrial proteins from a clinical *E. coli* strain CI6. Identical results were obtained in experiments using polyclonal anti-S-FimH (1-25) instead of monoclonal antibody IIF4.

and one of *Citrobacter freundii* as previously described^{15,16}. Control Pap pili were isolated from a strain of uropathogenic *E. coli*¹³. As probes, we used a polyclonal rabbit antiserum raised against a synthetic peptide S-FimH (1-25) copying the N-terminal 25 residues of the FimH tip protein, a mouse IgG2A monoclonal antibody, IIF4, raised against an SDS-gel-purified preparation of FimH, and a 60-base synthetic oligonucleotide copying the 5' end of the mature FimH structural gene sequence⁹. The FimH tip protein was isolated from *E. coli* ORN103 (kindly provided by Dr Paul Orndorff) transformed with a mutant plasmid pUT2004 derived from pSH2, which contains an 11.2 kb

fragment of *E. coli* chromosomal DNA encoding all the genes necessary for the expression of functional type 1 fimbriae^{7,10}. The procedures for isolating FimH and for preparation and characterization of the monoclonal anti-FimH antibodies will be described elsewhere (D.S., E.H.B. and S.N.A., manuscript in preparation).

The enzyme linked immunosorbent assay titres of the monoclonal antibody against FimH were 1:818,200 against isolated type 1 fimbriae and 1:6,400 against S-FimH (1-25); those of the polyclonal antibodies were 1:2,500 and 1:102,400, respectively. Both reacted in western blots with a protein band of M_r 28-31 K in each of the genera and species of type 1 fimbriae tested (Fig. 1). Furthermore, both antibodies inhibited the epithelial cell binding of each of the bacterial species tested except *E. coli* CI6 which expresses Pap but not type 1 fimbriae (Table 1).

Binding of the anti-*E. coli* FimH antibodies to the different enterobacterial species was probed with gold-conjugated protein A¹⁰. The gold particles decorated the type 1 fimbriae of each of the bacterial species tested at sites identical to those previously reported for *E. coli* type 1 fimbriae¹⁰, namely at the tips and at long intervals along the length of the fimbriae (Fig. 2). The binding of the monoclonal antibody was not inhibited by α -methyl-D-mannoside, suggesting that it is not directed against the D-mannose-binding site of FimH *per se*.

Finally, the radiolabelled¹⁷ oligonucleotide probe corresponding to the 5' end of *fimH* hybridized under high stringency conditions (less than 1% mismatch) in the dot blots¹⁸ with the total genomic DNA of all strains tested except type 1 fimbriated *E. coli* CI5 and *K. pneumoniae* 1376, and control Pap fimbriated *E. coli*. These results indicate a high degree of conservation of the D-mannose binding moiety both at the protein and nucleotide-sequence levels.

As it has become apparent that adhesion to mucosal surfaces is a prerequisite for the establishment of bacterial infections¹⁹, efforts have focused on finding ways to prevent bacterial adherence and colonization. Type 1 fimbrial vaccines have been tested with limited success in animals and man^{20,21}. One of the problems has been the high degree of heterogeneity of the major structural subunit¹¹; antibodies raised against the FimA subunit of one type 1 fimbrial species cross-react with only a limited number of those of other species. Furthermore, antibodies directed exclusively toward FimA often fail to block adhesion even of the homologous strains. In contrast, antibodies directed toward FimH, the adhesion protein of type 1 fimbriae, are

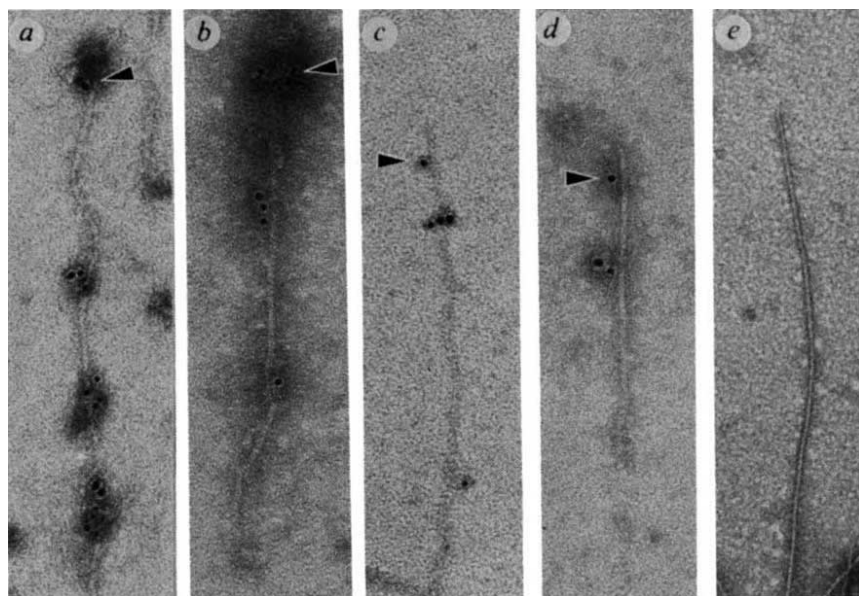


Fig. 2 Electron micrographs of various enterobacterial fimbrial species reacted with anti-FimH antibodies followed by gold-conjugated protein A¹⁰. Monoclonal antibody IIF4 was reacted with a, *E. coli* 1,400; b, *K. pneumoniae*; c, *S. marcescens* CI 1,495; d, *C. freundii* 39; and e, control Pap fimbriated *E. coli* CI6. In each case, the isolated fimbriae were applied to Formvar-coated copper grids, reacted with a 1:5,000 dilution of antibody IIF4 for 5 min, washed several times, exposed to gold (10 nm diameter)-conjugated protein A, washed again and negatively stained with phosphotungstic acid as described¹⁰. Identical results were obtained when the monoclonal antibody was substituted with polyclonal anti-S-FimH (1-25) (data not shown). Magnification, $\times 122,000$.

Table 1 Inhibition of the adhesion of type 1 fimbriated Enterobacterial species to human epithelial cells by antibodies

Species of Enterobacteriaceae	Adhesion of bacteria to epithelial cells in the presence of:	
	IIF4	Anti-S-FimH (1-25)
<i>E. coli</i> C15	14.4	15.1
<i>E. coli</i> C1 1400	18.1	23.3
<i>K. pneumoniae</i> 1,376	36.2	32.7
<i>K. pneumoniae</i> 59	42.9	51.5
<i>S. marcescens</i> 1,495	19.1	36.7
<i>S. marcescens</i> 50	26.5	37.5
<i>C. freundii</i> 39	37.4	43.7
<i>E. coli</i> C16	98.0	96.8
(Pap fimbriated)		

In this assay, 10^8 microorganisms were preincubated with a 1:100 dilution of monoclonal antibody IIF4, control ascites fluid, pre-immune rabbit serum, or rabbit polyclonal anti-S-FimH (1-25) before mixing with 10^6 human buccal epithelial cells. The dilutions of antibodies used failed to agglutinate any of the bacteria tested as determined by visual examination of antibody-bacterial mixtures by phase-contrast microscopy. Similar results were obtained with Fab fragments prepared from anti-S-FimH (1-25). Pre-immune sera and unrelated monoclonal antibodies had no effect (>95% of control values obtained in the absence of serum) on adhesion of any of the organisms tested. All non-adhesive organisms were removed by differential centrifugation and the number of adhering bacteria was determined by light microscopy as described¹⁵. The adhesion of bacteria is expressed as a percentage of control: (the number of adherent bacteria per 50 mucosal cells in the presence of antibodies/number of adherent bacteria per 50 mucosal cells in the absence of antibodies) $\times$ 100.

broadly cross-reactive not only with type 1 fimbriae of heterologous *E. coli* strains¹¹ but also, as shown here, with those of heterologous enterobacterial genera and species. Whether the broadly antiadhesive FimH vaccine will prove to be protective *in vivo* against infections due to various virulent genera and species of the family Enterobacteriaceae bearing type 1 fimbriae remains to be determined.

We thank Dr Roberta Sumrada for synthesizing the oligonucleotide probe in The Molecular Resource Center, Department of Microbiology, University of Tennessee, Mrs Loretta Hatmaker, Carolyn Costlow and Edna Chiang for technical assistance and Mrs Johnnie Smith and Sara Brewer for secretarial assistance. This work was supported by research funds from the US Veterans Administration and by research grants to S.A. and E.B. from the US Public Health Service.

Received 22 August; accepted 7 November 1988.

- Brinton, C. C. Jr *Trans. N.Y. Acad. Sci.* **27**, 1003-1054 (1965).
- Ofek, I., Mirelman, D. & Sharon, N. *Nature* **265**, 623-625 (1977).
- Duguid, J. P. & Old, D. C. in *Bacterial Adherence, Receptors and Recognition* (ed. Beachey, E. H.) 184-217 (Chapman and Hall, London 1980).
- Salit, J. E. & Gotschlich, E. C. *J. exp. Med.* **146**, 1169-1181 (1977).
- Eshdat, Y., Silverblatt, F. J. & Sharon, N. *J. Bact.* **148**, 308-314 (1981).
- Mauer, L. & Orndorff, P. E. *FEMS Microbiol. Lett.* **30**, 59-66 (1985).
- Minion, F. G., Abraham, S. N., Beachey, E. H. & Goguen, J. J. *Bact.* **165**, 1033-1036 (1986).
- Mauer, L. & Orndorff, P. E. *J. Bact.* **169**, 640-645 (1987).
- Klemm, P. & Christiansen, G. *Molec. gen. Genet.* **208**, 439-455 (1987).
- Abraham, S. N., Goguen, J., Sun, D., Klemm, P. & Beachey, E. H. *J. Bact.* **169**, 5530-5536 (1987).
- Hanson, M. S. & Brinton, C. C. Jr *Nature* **332**, 265-268 (1988).
- Giampapa, C. S., Abraham, S. N., Chiang, T. M. & Beachey, E. H. *J. biol. Chem.* **263**, 5362-5367 (1988).
- Lindberg, F., Lund, B., Johansson, L. & Normark, S. *Nature* **328**, 84-87 (1987).
- Moch, T., Hoschutsky, H., Hacker, J., Kronck, K. D. & Jann, K. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3462-3466 (1987).
- Abraham, S. N., Hasty, D. L., Simpson, W. A. & Beachey, E. H. *J. exp. Med.* **158**, 1114-1128 (1983).
- Dodd, D. C. & Eisenstein, B. I. *Infect. Immunity* **38**, 764-773 (1982).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Buchanan, B. et al. *J. Bact.* **162**, 799-803 (1985).
- Beachey, E. H. *J. infect. Dis.* **143**, 325-345 (1981).
- Silverblatt, F. J. & Cohen, L. S. *J. clin. Invest.* **64**, 333-336 (1979).
- Levine, M. M. et al. *Microbiol. Rev.* **47**, 526-528 (1983).

Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells

R. Lindsay Williams*, Douglas J. Hilton†, Shirley Pease*, Tracy A. Willson†, Colin L. Stewart*||, David P. Gearing†, Erwin F. Wagner*||, Donald Metcalfe†, Nicos A. Nicola† & Nicholas M. Gough†

* European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, FRG

† The Walter and Eliza Hall Institute of Medical Research, Post Office, Royal Melbourne Hospital, Victoria 3050, Australia

Embryonic stem (ES) cells, the totipotent outgrowths of blastocysts^{1,2}, can be cultured and manipulated *in vitro* and then returned to the embryonic environment where they develop normally and can contribute to all cell lineages³⁻⁹. Maintenance of the stem-cell phenotype *in vitro* requires the presence of a feeder layer of fibroblasts^{1,2,10} or of a soluble factor, differentiation inhibitory activity (DIA) produced by a number of sources^{5,11,12}; in the absence of DIA the ES cells differentiate into a wide variety of cell types. We recently noted several similarities between partially purified DIA and a haemopoietic regulator, myeloid leukaemia inhibitory factor (LIF), a molecule which induces differentiation in M1 myeloid leukaemic cells and which we have recently purified, cloned and characterized¹³⁻¹⁸. We demonstrate here that purified, recombinant LIF can substitute for DIA in the maintenance of totipotent ES cell lines that retain the potential to form chimaeric mice.

Initially, representative ES and embryonic carcinoma (EC) cell lines were assayed for LIF receptors. M1 cells and normal murine monocytes display a small number (50-200) of high-affinity LIF receptors with an apparent dissociation constant of 50-150 pM, whereas a wide variety of other normal or neoplastic haemopoietic cells do not display LIF receptors¹⁹. Scatchard analyses of ¹²⁵I-labelled LIF binding to the totipotent ES cell line EKcs-1 (ref. 20) and the EC cell lines PCC3A1 and F9 (refs 21, 22) reveal that these cells express 295, 190 and 330 high-affinity LIF receptors per cell with apparent dissociation constants of approximately 90 pM (Fig. 1a). All other ES and EC lines tested—D3 (refs 9, 23); NG2 (ref. 24); PC13 and P19 (ref. 22)—bound similar levels of LIF (data not shown). The binding of ¹²⁵I-labelled LIF to EKcs-1, PCC3A1 and M1 cells was specific, as it could be quantitatively competed with purified native LIF or with purified recombinant LIF derived from yeast or *Escherichia coli*. No competition was observed with a range of other factors, including granulocyte colony stimulating factor, granulocyte/macrophage colony stimulating factor, interleukin 6, acidic and basic fibroblast growth factor, nerve growth factor, transforming growth factor β and insulin-like growth factors I and II (data not shown). Autoradiography (Fig. 1b, c) of F9 cells after ¹²⁵I-labelled LIF binding showed that essentially all (96%) of the cells expressed LIF receptors, although the varying grain counts indicated some heterogeneity in receptor numbers, similar to that found with other cell types¹⁹.

To test whether LIF could maintain the stem-cell phenotype of ES cells, four independently derived ES cell lines D3, EKcs-1, CBL63 (ref. 23) and HD5 (C.L.S., unpublished results) were propagated in highly purified, recombinant LIF (see Fig. 2 legend). At concentrations of LIF of 1,000-5,000 units ml⁻¹, as in the presence of feeder cells or medium conditioned by the human bladder carcinoma cell line 5637, more than 95% of colonies displayed the stem-cell phenotype of compact colonies

|| Present addresses: Roche Institute for Molecular Biology, Nutley, New Jersey (C.L.S.); Institute for Molecular Pathology, Dr Bohr-Gasse 7, 1030 Vienna, Austria (E.F.W.).

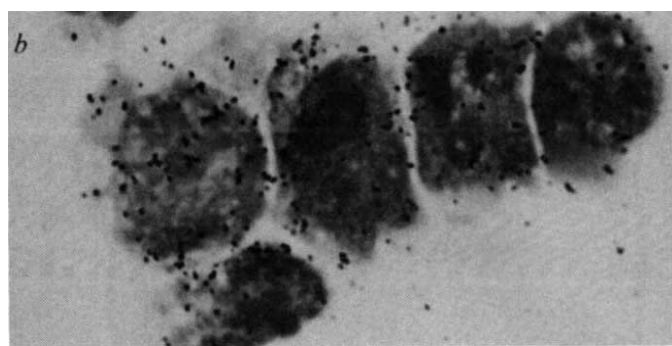
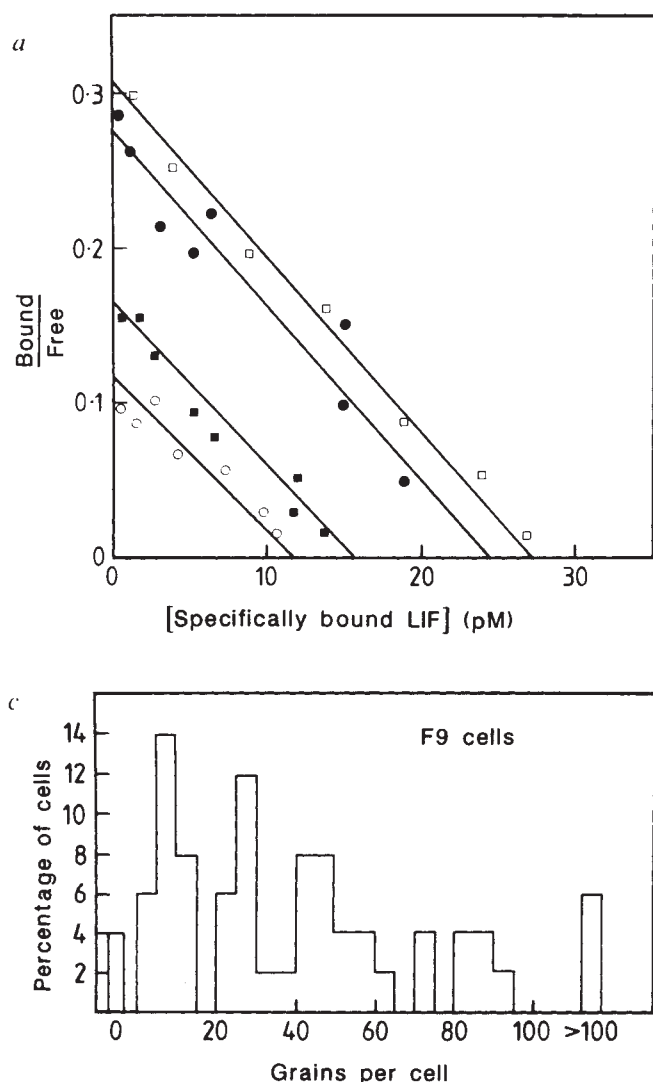


Fig. 1 LIF receptors on ES and EC cells. *a*, Scatchard analysis of ^{125}I -labelled LIF binding to F9 ($\square$), EKcs-1 ($\bullet$), PCC3A-1 ($\blacksquare$) and M1 ($\circ$) cells. Saturation curves for binding were analysed by the method of Scatchard by plotting the amount of LIF specifically bound (defined as the difference between binding observed in the absence and presence of excess unlabelled LIF) versus the ratio of bound to free LIF. Free LIF values were adjusted for the per cent of ^{125}I -labelled LIF capable of binding specifically to LIF receptors, in this experiment determined to be 75%. The apparent dissociation constant for the interaction of LIF with its receptor was determined from the slopes of the curves and the receptor number from their intercepts with the ordinate. Results in *a* were standardized to 5×10^6 cells per point and the mean of duplicate points are shown and curves were fitted using the Ligand program²⁸. *b*, Autoradiography of F9 EC cells labelled with ^{125}I -labelled LIF. *c*, Quantitation of silver grains on F9 EC cells after binding of ^{125}I -labelled LIF.

Methods. Purified native murine LIF¹⁷ was iodinated and binding was carried out as described¹⁹. Autoradiography was performed as described²⁹ on glutaraldehyde-fixed cytocentrifuge smears. Labelled cells were exposed to Kodak NBT2 photographic emulsion for 32 days, developed, fixed and stained with 10% v/v Giemsa solution. Photomicrographs were taken under oil ($\times 700$). Grain counts were scored on 200 consecutive cells. Background grain counts determined from adjacent cell-free areas varied from 0–4 grains and were subtracted from cellular grain counts. Parallel autoradiographs performed on cells incubated in the presence of a 50-fold excess of unlabelled LIF showed no significant grains on cells above background.

of small cells with a large nuclear to cytoplasmic ratio (Figs 2 and 3). In contrast, ES cells maintained in normal culture medium differentiated over a period of 3–6 days and the resultant colonies contained large, flat differentiated cells. The proportion of colonies having a stem-cell phenotype was related to the concentration of LIF in the culture (Fig. 2). Consistent with the similar binding affinities of LIF receptors on M1 and ES cells, the dose-response curve for the action of LIF on ES cells was similar to that on M1 cells, with 50% maximal activity occurring

Table 1 Production of chimaeric mice from D3 ES cells maintained in LIF

Number of blastocysts transferred	Number of progeny		Number of chimaeras	
	Embryos	Pups	Embryos	Pups
39	22 (56%)	–	15 (68%)	–
101	–	37 (37%)	–	25 (67%)

D3 ES cells, homozygous for the glucose phosphate isomerase I^a allele, were maintained in 1,000–5,000 units ml^{-1} H-LIF for 7–22 passages (32 to 100 cell generations) and then injected into ICR or C57BL/6J blastocysts⁷ homozygous for the glucose phosphate isomerase I^b allele. The blastocysts were then transferred into pseudo-pregnant foster mothers and the resulting mid-gestation embryos or mice tested for chimaerism by analysis of the glucose phosphate isomerase I isoenzymes or coat pigmentation³.

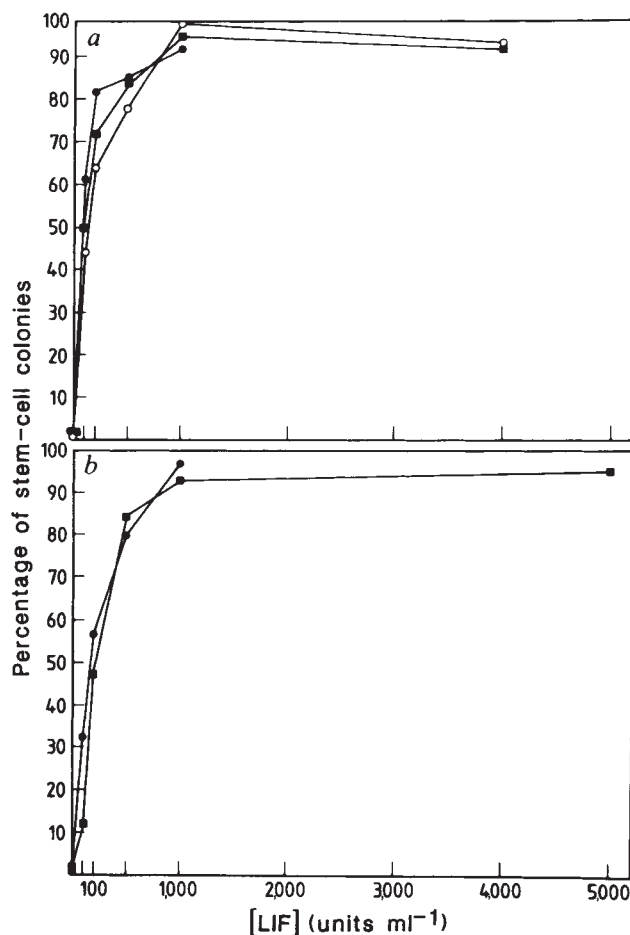
at around 50–100 units ml^{-1} (ref. 16; Fig. 2). These results are also consistent with the concentration of LIF in 5637 conditioned medium (2,000–5,000 units ml^{-1}) and in conditioned medium from primary embryonic fibroblasts used as feeder layers for ES cells (500–1,000 units ml^{-1}). Incubation of ES cells in the presence of LIF from 0–5,000 units ml^{-1} did not influence the number or size of colonies formed after six days in culture ruling out the possibility that LIF was selecting for a specific subpopulation of cells.

Long-term maintenance of the four ES cell lines in LIF for up to 22 passages (approximately 100 cell generations) did not alter the growth characteristics of these cells as their dose dependency on LIF was similar to that seen with ES cell lines explanted directly from a feeder layer or from 5637 conditioned medium (Fig. 2). The stem-cell phenotype of ES cells cultured for multiple passages in the presence of LIF was confirmed by immunofluorescence with the ECMA-7 antibody which recognizes a cell-surface stem-cell-specific antigen²⁵: ES cells cultured in the presence of LIF expressed the stem-cell marker, whereas in the absence of LIF less than 1% did so (Fig. 3e, f).

The ability of ES cells maintained in LIF for multiple passages to differentiate spontaneously into diverse cell types in culture suggests that the cells remain totipotent. To confirm their developmental potential, D3 ES cells maintained in LIF for 7–22 passages were injected into blastocysts that were then introduced into pseudo-pregnant females. Over 65% of the progeny analysed contained tissues derived from the injected

Fig. 2 Growth of ES cells in recombinant LIF. *a*, HD5 cells previously maintained in 80% 5637 conditioned medium for eight passages were transferred to culture media containing 0–5,000 units ml^{-1} of purified, recombinant yeast-derived human LIF (H-LIF; see below) ($\square$ — $\square$) or purified, recombinant *E. coli*-derived mouse LIF (M-LIF; see below) ($\circ$ — $\circ$). HD5 cells maintained in medium containing 1,000 units ml^{-1} H-LIF for a further 13 passages were then transferred to 0–1,000 units ml^{-1} M-LIF ($\bullet$ — $\bullet$). *b*, D3 cells maintained on mouse embryo fibroblasts for 10 passages were transferred to media containing 1,000–5,000 units ml^{-1} H-LIF and after a further 7 or 15 passages the cells were transferred into media containing 0–5,000 units ml^{-1} of H-LIF ($\blacksquare$ — $\blacksquare$) or 0–1,000 units ml^{-1} M-LIF ($\bullet$ — $\bullet$) respectively.

Methods. Two versions of recombinant LIF were used throughout these experiments with similar results: human LIF produced in yeast cells¹⁴ (H-LIF) and mouse LIF produced in *E. coli* using the pGEX expression system³⁰ (M-LIF). The yeast-derived LIF was purified by sequential chromatography on CM Sepharose CL6B, DEAE Sepharose CL6B and Sepharose S300, similarly to native LIF¹⁷. The *E. coli*-derived LIF was purified by glutathione-agarose affinity chromatography³⁰ and C8 reverse-phase HPLC. Both preparations of LIF were greater than 95% pure as assessed SDS PAGE. The cross-reactivity of human LIF on murine ES cells is consistent with its cross-reactivity on other murine cells¹⁴. LIF activity was based on the ability to induce M1 differentiation, defining 50 units ml^{-1} as the concentration of LIF that induces 50% of M1 colonies to display the differentiated phenotype¹⁶. The specific activity of native LIF is 1.2×10^8 units per mg ¹⁷. Culture conditions and media used for the growth of ES cells including the preparation of conditioned media have been described elsewhere^{7,11}. To assess the maintenance of the stem-cell phenotype, a single-cell suspension (10^3 cells in 35-mm tissue culture dishes pre-coated with a 0.1% gelatin solution; duplicate cultures for each concentration of LIF) was plated in media containing 0–5,000 units ml^{-1} purified, recombinant LIF. After seven days, colonies were stained and the ratio of stem-cell and differentiated colonies determined. Over 90% of the colonies examined consisted of distinct stem or differentiated cell populations; the few colonies consisting of mixed cell populations were characterized according to the relative proportions of each cell type.



cells (Table 1), with levels of overt chimaerism as high as 90% in individual mice. Furthermore, analysis of the organs of four chimaeras confirmed that ES cells maintained in LIF could contribute extensively to the development of all of the somatic tissues (Table 2).

Our observation that LIF retards the differentiation of ES cells is intriguing in view of its identification within the haemopoietic system as a molecule which induces the differentiation of M1 cells. These divergent actions of LIF occur at the same concentration of factor and underscore the notion that the response elicited by the binding of a protein regulator to a unique cell-surface receptor is determined ultimately by the set of intracellular signals activated, and suggest that M1 cells and ES cells express distinct intracellular signalling pathways with which the LIF receptor can interact.

There are a number of biochemical similarities between DIA and LIF: both have a relative molecular mass of about 50,000 and are heavily glycosylated (refs 12, 13, 15, 17, 31) and both are produced by the human bladder carcinoma cell line 5637. Moreover, it has recently been demonstrated that purified DIA from Buffalo rat liver cells, like LIF, stimulates the proliferation of murine DA1.1a cells³¹, which together with our results suggest that DIA and LIF are indeed the same molecule and that this haemopoietic regulator may also have a role in early embryogenesis.

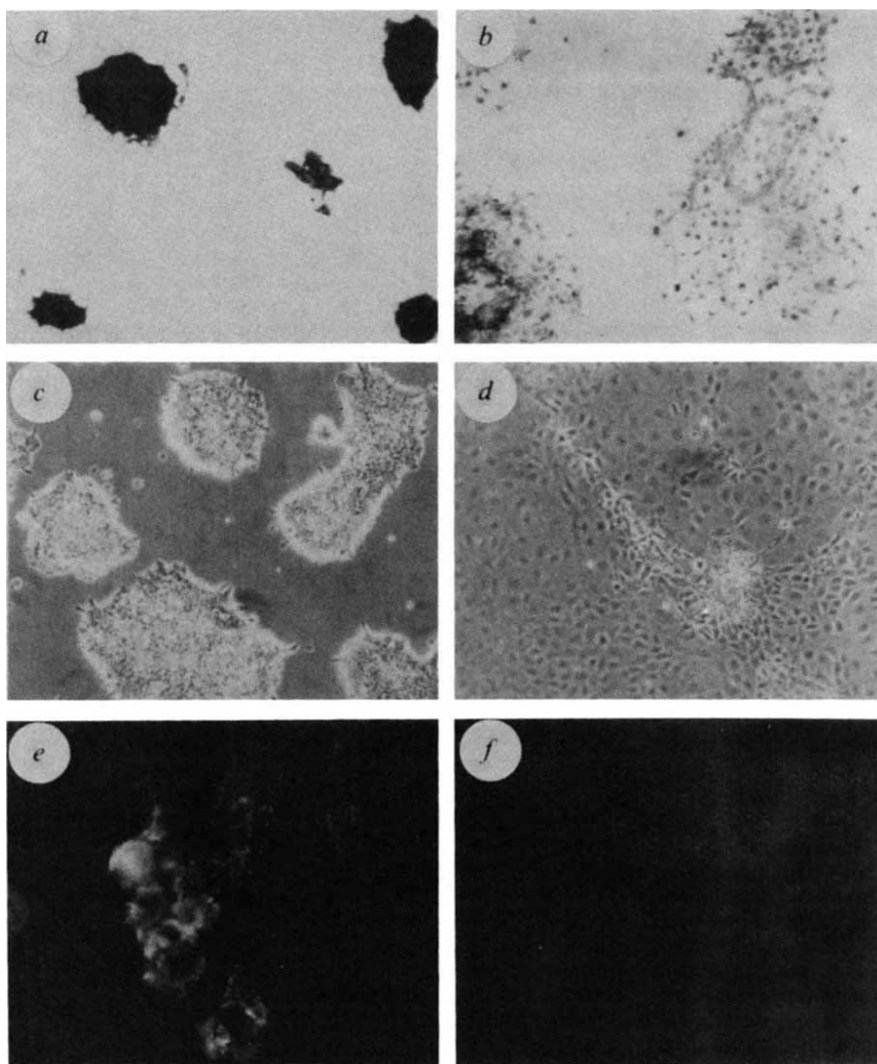
It should now be possible to study embryos directly to determine the pattern of expression of LIF and its receptor during embryogenesis and to investigate its role in differentiation. The availability of purified, recombinant LIF will assist the *in vitro* analysis of the differentiation of ES cells and should facilitate

Table 2 Percentage tissue contributions in individual D3 chimaeric mice

Chimaera	Necropsy age	C	Bl	Sp	P	Li	T	H	Lu	G	K	M	B	Sa
D3-1	13d	35	0	35	20	10	20	40	30	10	35	30	35	20
D3-2	14d	40	15	35	30	45	30	50	35	20	30	50	50	25
D3-3	11d	90	50	50	35	50	40	60	45	50	50	70	50	55
D3-4	11d	50	50	50	30	40	40	50	50	35	50	50	20	30

Strain-specific markers in the coat were at the A (*agouti*) locus; other tissue analysis was based on glucose phosphate isomerase isoenzyme strain differences (see Table 1). The extent of tissue chimaerism was determined in two D3-ICR (numbers 1 and 2) and two D3-C57BL/6J chimaeras (numbers 3 and 4). Tissues analysed: C, coat; Bl, Blood; Sp, spleen; P, pancreas; Li, liver; T, thymus; H, heart; Lu, lungs; G, gonads; K, kidneys; M, muscle; B, brain; Sa, salivary gland.

Fig. 3 ES cell morphology in the presence of recombinant LIF. HD5 ES cells cultured in the presence of 80% 5637 conditioned medium were assayed for the ability of purified recombinant LIF to maintain the stem-cell phenotype by transfer to media containing 1,000 units ml^{-1} M-LIF (a), or to normal culture media (b) as described in Fig. 2. After seven days the colonies were stained with Giemsa. Compact stem-cell colonies could be distinguished from diffuse differentiated colonies. D3 cells maintained in H-LIF for 15 passages were assayed for the ability to differentiate by transfer into media containing 1,000 units ml^{-1} M-LIF (c) or normal culture media (d) as described in Fig. 2. Immunofluorescence of the cells in the two D3 colony types was carried out using the ECMA-7 monoclonal antibody²⁵ which recognizes a stem cell-specific cell-surface antigen. Cell-surface-specific immunofluorescence was detected on over 90% of the cells maintained in media containing 1,000 units ml^{-1} recombinant LIF (e) but on less than 1% of the cells maintained in normal culture media (f). The field of view shown in f contains 21 cells. Immunofluorescence was performed as described²².



the generation and culture of these cells which, with the advent of gene targeting experiments^{26,27} are an important alternative route for the generation of transgenic mice.

We thank Rolf Kemler for the D3 and CBL63 cell lines and ECMA-7 antibody, David Stoddart for maintaining the mouse colony and Austin Smith, John Heath, Dick Cotton, Jerry Adams and Ann Johnsson for useful discussions. The work at

WEHI was supported by the NH and MRC (Canberra), the Anti-Cancer Council of Victoria and Australian Medical Research and Development Corporation and the work at EMBL was partly supported by the Deutsche Forschungsgemeinschaft. *Note added in proof:* D3 ES cells maintained in LIF for 16 passages *in vitro* have been transmitted through the germ-line of the first male D3-C57BL/6J chimaera tested.

Received 18 August; accepted 19 October 1988.

- Evans, M. J. & Kaufman, M. H. *Nature* **292**, 154-156 (1981).
- Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7634-7638 (1981).
- Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. *Nature* **309**, 255-256 (1984).
- Robertson, E. J., Bradley, A., Kuehn, M. & Evans, M. *Nature* **323**, 445-448 (1986).
- Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. *Nature* **326**, 292-295 (1987).
- Kuehn, M. R., Bradley, A., Robertson, E. J. & Evans, M. J. *Nature* **326**, 295-298 (1987).
- Williams, R. L., Courtneidge, S. A. & Wagner, E. F. *Cell* **52**, 121-131 (1988).
- Stewart, C. L., Vanek, M. & Wagner, E. F. *EMBO J.* **4**, 383-388 (1985).
- Gossler, A., Doetschmann, T., Korn, R., Serfling, E. & Kemler, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9065-9069 (1986).
- Martin, G. R. & Evans, M. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1441-1445 (1975).
- Smith, A. G. & Hooper, M. L. *Dev. Biol.* **121**, 1-9 (1987).
- Koopman, P. & Cotton, R. G. H. *Exptl Cell Res.* **154**, 233-242 (1984).
- Gearing, D. P. *et al. EMBO J.* **6**, 3995-4002 (1987).
- Gough, N. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 2623-2627 (1988).
- Hilton, D. J., Nicola, N. A., Gough, N. M. & Metcalf, D. *J. biol. Chem.* **263**, 9238-9443 (1988).
- Metcalf, D., Hilton, D. J. & Nicola, N. A. *Leukaemia* **2**, 216-221 (1988).
- Hilton, D. J., Nicola, N. A. & Metcalf, D. *Analyt. Biochem.* **173**, 359-367 (1988).

- Simpson, R. J. *et al. Eur. J. Biochem.* **175**, 541-547 (1988).
- Hilton, D. J., Nicola, N. A. & Metcalf, D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5971-5975 (1988).
- Wagner, E. F., Keller, G., Gilboa, E., Rüther, U. & Stewart, C. *Cold Spring Harb. Symp. quant. Biol.* **50**, 691-700 (1985).
- Jakob, J., Bonn, T., Gaillard, J., Nicolas, J. F. & Jacob, F. *Ann. Microbiol. (Inst. Pasteur)* **124B**, 269-282 (1973).
- Rudnicki, M. A. & McBurney, M. W. in *Teratocarcinomas and Embryonic Stem Cells, a Practical Approach* (ed. E. J. Robertson) 19-49 (IRL, Oxford, 1987).
- Doetschman, T., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. *J. Embryol. exp. Morphol.* **87**, 27-45 (1985).
- Dewey, M. J., Martin, D. W., Martin, G. R. & Mintz, B. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5564-5568 (1977).
- Kemler, R. in *Progress in Developmental Biology Band 26* (ed. Sauer, H. W.) 175 (Fischer, Stuttgart, 1980).
- Thomas, K. R. & Capecchi, M. R. *Cell* **51**, 503-512 (1987).
- Doetschman, T. *et al. Nature* **330**, 576-577 (1987).
- Munson, P. J. & Rodbard, D. *Analyt. Biochem.* **107**, 220-237 (1980).
- Nicola, N. A. & Metcalf, D. *J. cell. Physiol.* **124**, 313-321 (1985).
- Smith, D. B. & Johnson, K. S. *Gene* **67**, 31-40 (1988).
- Smith, A. G. *et al. Nature* **336**, 688-690 (1988).

Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides

Austin G. Smith*, John K. Heath*†, Deborah D. Donaldson‡, Gordon G. Wong‡, J. Moreau§, Mark Stahl‡ & David Rogers‡

* Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

‡ Genetics Institute, 87 Cambridge Park Drive, Cambridge, Boston, Massachusetts 02140, USA

§ INSERM U211, 1 Rue Gaston-Veill, 44035 Nantes, Cedex France

Murine embryonic stem (ES) cells are pluripotent cell lines established directly from the early embryo^{1,2} which can contribute differentiated progeny to all adult tissues, including the germ-cell lineage³, after re-incorporation into the normal embryo. They provide both a cellular vector for the generation of transgenic animals⁴ and a useful system for the identification of polypeptide factors controlling differentiation processes in early development⁵. In particular, medium conditioned by Buffalo rat liver cells contains a polypeptide factor, ES cell differentiation inhibitory activity (DIA), which specifically suppresses the spontaneous differentiation of ES cells *in vitro*, thereby permitting their growth as homogeneous stem cell populations in the absence of heterologous feeder cells⁶. ES cell pluripotentiality, including the ability to give rise to functional gametes, is preserved after prolonged culture in Buffalo rat liver media as a source of DIA⁷. Here, we report that purified DIA is related in structure and function to the recently identified haemopoietic regulatory factors human interleukin for DA cells^{8,9} and leukaemia inhibitory factor¹⁰. DIA and human interleukin DA/leukaemia inhibitory factor have thus been identified as related multifunctional regulatory factors with distinct biological activities in both early embryonic and haemopoietic stem cell systems.

DIA purified from Buffalo rat liver (BRL)-conditioned medium is a single chain glycoprotein of relative molecular mass $M_r = 43,000$ formed by extensive glycosylation of a core polypeptide of apparent $M_r = 20,000$ (Fig. 1). Cells grown in the presence of 10 ng ml^{-1} purified DIA retain a typical ES cell morphology (Fig. 2a) and biochemical phenotype. Immunofluorescence on fixed cells indicates that the SSEA-1 cell surface antigen¹¹ is expressed by at least 78% of cells cultured in the presence of DIA. Suppression of differentiation can be observed at concentrations of 0.2 ng ml^{-1} DIA. ES cells have been propagated for over 30 generations in media supplemented with purified DIA without overt differentiation. Removal of DIA from the culture medium results in spontaneous differentiation of the ES cells, with a concomitant change in morphology (Fig. 2c) and cessation of SSEA-1 expression.

While investigating possible actions in non-ES cell systems, it was found that DIA was capable of supporting continued growth *in vitro* of DA-1a MoMuLV-induced leukaemia cells. In the accompanying paper¹², the isolation of a complementary DNA clone (pC10-6R) encoding a factor (human interleukin for DA cells, HILDA, ref. 9) with similar actions on DA-1a cells is reported. The sequence of pC10-6R is identical to that reported for human leukaemia inhibitory factor (LIF)¹³. The pC10-6R cDNA clone was cloned into the eukaryotic expression vector PMT2 and transfected into COS cells. ES cells grown in a 1/200 dilution of medium conditioned by COS cells transfected with the HILDA/LIF cDNA clone maintained typical ES cell morphology and biochemical phenotype. Induction of the characteristic change in morphology and loss of SSEA-1 expression

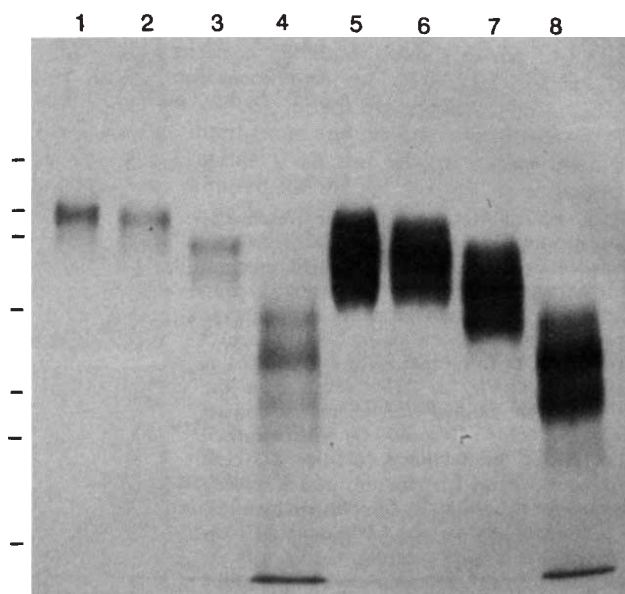


Fig. 1 N-glycanase digestion of purified radioiodinated DIA (lanes 1-4) and HILDA/LIF (lanes 5-7). Lanes 1 and 5 are untreated DIA and HILDA/LIF respectively. Lanes 2 and 6 are samples incubated in the absence of enzyme. Lanes 3 and 7 are digests performed under non-denaturing conditions, and lanes 4 and 8 are digests of denatured protein. Relative molecular mass markers are $M_r = 66,000, 45,000, 36,000, 29,000, 24,000, 20,100$ and $14,200$.

Methods. DIA and HILDA/LIF were purified by methods to be described elsewhere (manuscript in preparation) from BRL-conditioned medium and medium conditioned by COS-7 cells transfected with the pC10-6R HILDA/LIF cDNA clone. Both preparations gave single symmetrical peaks on microbore reverse-phase chromatography. Purity was estimated as $>95\%$ judged by SDS-PAGE and silver staining. Purified proteins were quantitated by amino-acid analysis. The amino-acid composition of purified HILDA/LIF was in agreement with that predicted from the sequence of the pC10-6R cDNA clone. The recovery of HILDA/LIF from COS cell conditioned medium was $170 \mu\text{g litre}^{-1}$ and DIA was recovered from BRL-conditioned medium at approximately $200 \text{ ng litre}^{-1}$. DIA and HILDA/LIF were iodinated by incubation of 500 ng of protein with $1 \text{ mCi } ^{125}\text{I}$ iodide in the presence of immobilized lactoperoxidase/glucose oxidase (Enzymobeads, Biorad) and $1\% \beta$ -D-glucose in 0.1 M phosphate buffer $\text{pH } 7.2$ for 60 min at 25°C . Specific activities were $7.5 \times 10^3 \text{ c.p.m. ng}^{-1}$ (DIA) and $2 \times 10^4 \text{ c.p.m. ng}^{-1}$ (HILDA/LIF). N-glycanase digestions were performed for 18 h at 37°C in 0.2 M sodium phosphate/ $1.25\% \text{ NP-40}$, $\text{pH } 8.6$, with 10 units ml^{-1} of N-glycanase (Genzyme). Samples were denatured by boiling for 10 min in $0.5\% \text{ SDS}$, 100 mM mercaptoethanol. Digestions were terminated by addition of an equal volume of double strength Laemmli sample buffer, boiled for 5 min and applied to a 14% discontinuous SDS-PAGE gel¹⁵. After electrophoresis, the gel was dried and exposed FUJI-RX X-ray film for autoradiography.

was evident upon removal of conditioned medium (Fig. 2c). No activity was found in conditioned media from mock-transfected COS cells. ES cells could be maintained in the presence of HILDA/LIF for prolonged periods without biochemical or morphological differentiation; a minimum of 76% of the cells expressed the SSEA-1 antigen after culture for over 20 generations in pC10-6R transfected COS cell conditioned medium.

The activity present in medium conditioned by COS cells transfected with pC10-6R was purified to homogeneity by a modification of methods developed for the purification of BRL-derived DIA, monitoring by the ability to inhibit ES cell differentiation. HILDA/LIF purified from pC10-6R transfected COS cells is a single-chain molecule with multiple forms ranging from $M_r = 45,000$ to $M_r = 35,000$ (Fig. 1). Treatment of COS

† To whom correspondence should be addressed.

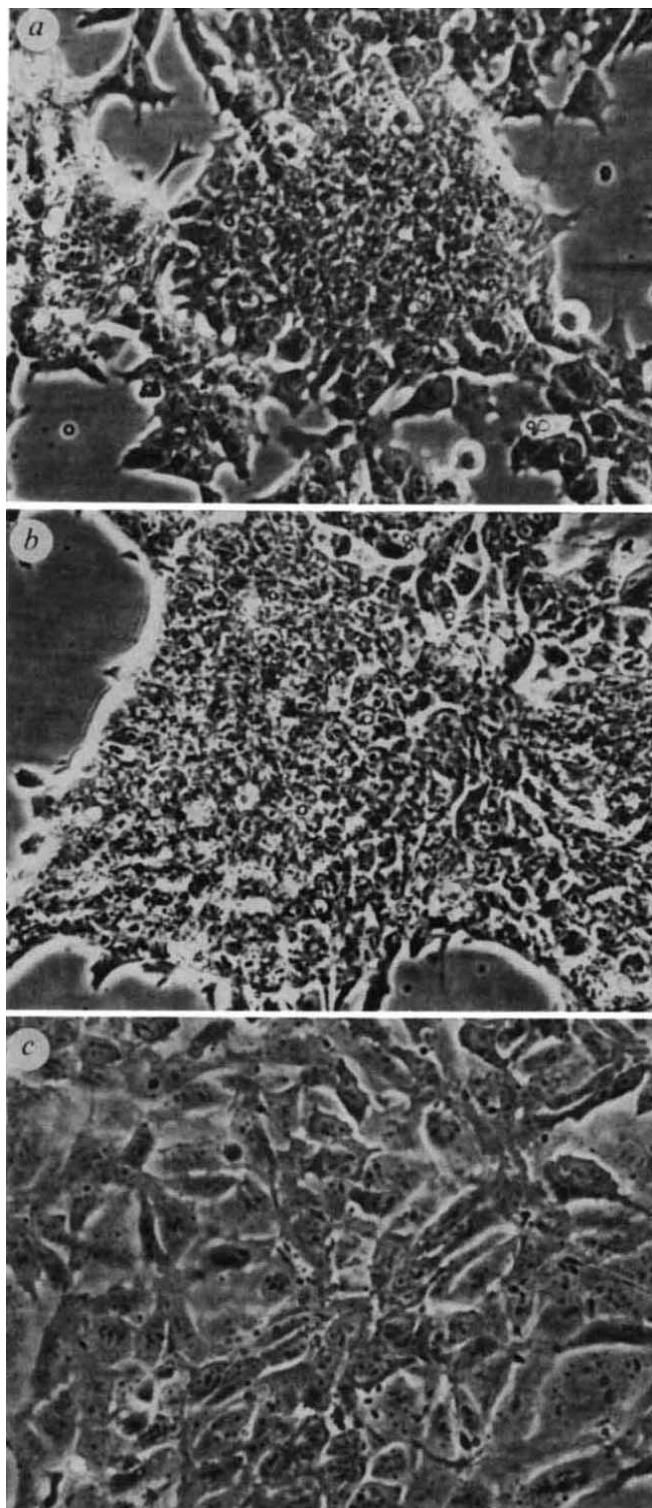


Fig. 2 Morphology of CP1 ES cells grown for six days in the presence of 10 ng ml^{-1} purified DIA (a), 10 ng ml^{-1} purified HILDA/LIF (b), and no additions (c). Cells were propagated on gelatinized tissue culture plates in the presence of Hams F12/DME (50:50), 10^{-4} M 2-mercaptoethanol, 20% (by volume) fetal calf serum (Sera-Lab, UK). ES cell stocks were routinely maintained in medium supplemented with 100-fold concentrated BRL conditioned medium to a final concentration of 1.5%. Final magnification $\times 225$.

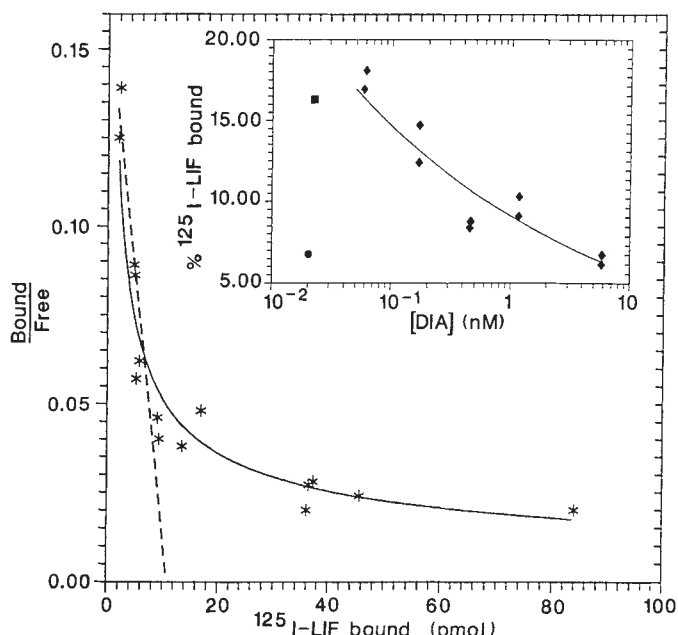


Fig. 3 Binding of ^{125}I -HILDA/LIF to CP1 ES cells. HILDA/LIF was iodinated as in the Fig. 1 legend. The biological activity of HILDA/LIF before and after the iodination reaction was equivalent. Inset: inhibition of ^{125}I -labelled HILDA/LIF binding to CP1 ES cells by purified DIA. CP1 ES cells were exposed to varying concentrations of purified DIA ($\blacklozenge$), or 250 ng purified HILDA/LIF ($\bullet$) or control media ($\blacksquare$). ^{125}I -labelled HILDA/LIF binding was determined using 1 ng of labelled ligand in each reaction.

Methods. Binding reactions were performed on monolayers of CP1 ES cells in binding buffer HAMS F12:DME (Dulbecco's modified Eagles medium) (50:50), 0.2% bovine serum albumin, 25 mM HEPES, pH 7.2 at 14°C for 2.5 hours. Monolayers were washed three times with ice-cold binding buffer at the end of the incubation and cell-associated radioactivity measured by solubilizing the cells in 1% NP-40 followed by gamma counting. Cell numbers (3×10^6 cells per 35-mm well) were determined from duplicate plates at the end of the incubation. Binding data was analysed as described¹⁶, assuming a relative molecular mass of 43,000 for DIA and HILDA/LIF. This method yields two apparent components; one corresponding to high affinity receptor sites and a second corresponding to low affinity/high capacity ('non-specific') sites. The high-affinity sites were considered to represent physiological HILDA/LIF receptors.

cell-derived HILDA/LIF with N-glycanase releases a series of lower relative molecular mass forms, the smallest of which has an apparent M_r of 20,000 (Fig. 1) in agreement with the predicted size of the polypeptide encoded by the pC10-6R cDNA¹². ES cells cultured in the presence of 10 ng ml^{-1} purified HILDA/LIF retained characteristic stem cell morphology (Fig. 2b) and inhibition of differentiation could still be observed at a concentration of 0.2 ng ml^{-1} . Williams and co-workers¹⁴ have independently observed that recombinant LIF has ES cell differentiation inhibiting properties.

The physical and functional similarities between DIA and HILDA/LIF, and the concentrations at which their biological effects are exerted, suggested that both agents might interact with related plasma membrane receptor systems. This was tested by examining the expression of specific cell-surface receptors on ES cells. ^{125}I -labelled HILDA/LIF binds to an apparently single class of high-affinity receptors on ES cells (dissociation constant $K_d = 80 \text{ pM}$, 4,500 sites per cell (Fig. 3). Binding of ^{125}I -HILDA/LIF is specifically inhibited in a concentration-dependent way by purified DIA (Fig. 3, inset). DIA and HILDA/LIF therefore interact with the same cell-surface receptor system expressed by murine ES cells.

The molecular characterization of polypeptide factors that specifically suppress the differentiation of pluripotent ES cells has several important consequences. First, DIA and HILDA/LIF may regulate the differentiation of pluripotential embryonic ectoderm during normal embryogenesis. Second, the underlying molecular mechanism of DIA/HILDA/LIF action is of interest as it controls a critical developmental decision. Interaction with specific cell-surface receptors may have an initiating role in this process. Finally, the availability of recombinant factor will facilitate the growth of pure ES cell populations *in vitro* for biochemical analysis and germ-line manipulation.

We are grateful to Peter Stern for antibodies, Tony Willis for amino-acid analyses, Liz Robertson for cells and Dylan Edwards, Anna-Maria Caccuri, Jenifer Nichols, Frances Bennett and Joanne Witek-Giannotti for assistance. We thank Lindsay Williams and Nick Gough for exchange of information and discussion. J.K.H. and A.G.S. are supported by the Cancer Research Campaign.

Note added in proof: 8/15 overtly chimaeric progeny have been obtained from two litters of mice derived from blastocysts injected with ES cells propagated in purified recombinant HILDA/LIF (J. Nichols, A.G.S. and J.K.H., in preparation).

Received 9 September; accepted 21 October 1988.

- Evans, M. J. & Kaufman, M. *Nature* **292**, 154–156 (1981).
- Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7634–7638 (1981).
- Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. *Nature* **309**, 255–256 (1984).
- Robertson, E. J. *Trends Genet.* **2**, 9–13 (1986).
- Heath, J. K., Smith, A. G., Wills, A. J. & Edwards J. *Cell Sci. Sup.* (in the press).
- Smith, A. G. & Hooper, M. L. *Dev. Biol.* **121**, 1–9 (1987).
- Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. *Nature* **326**, 292–2857 (1987).
- Moreau, J. F. *et al. Ann. Inst. Pasteur, Immunol.* **137**, 25–37 (1986).
- Moreau, J. F. *et al. J. Immun.* **138**, 3844–3849 (1987).
- Gearing, D. P. *et al. EMBO J.* **6**, 3995–4002 (1987).
- Solter, D. & Knowles, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5565–5569 (1978).
- Moreau, J-F. *et al. Nature* **336**, 690–692 (1988).
- Gough, N. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 2623–2627 (1988).
- Williams, R. L. *et al. Nature* **336**, 684–687 (1988).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Mendel, C. M. & Mendel, D. B. *Biochem. J.* **228**, 269–272 (1985).

Table 1 Levels of HILDA activity from different sources

Source	Activity (units ml ⁻¹)
Cell-line conditioned medium:	
C10-MJ2	< 10
C10-MJ2, induced	160
H23, induced	250
BRL 32	100
Oocyte translations:	
C10-MJ2, induced mRNA	30
C10-MJ2, superinduced mRNA	30
Transfected-cos-cell-conditioned medium:	
C10-MJ2 library primary pool #11	30
pC10-6R	36,000
pC10-6R(2)	36,000

HILDA activity was measured as described elsewhere⁵. One unit per ml is the inverse of the dilution at which the sample gave half of its maximal response. C10-MJ2 cells were induced with phytohaemagglutinin (PHA) and 12-*o*-tetradecanoylphorbol-13-acetate (TPA) as described elsewhere¹⁷. Conditioned medium from the C10-MJ2 cells was collected after 48 h and cells were collected for mRNA isolation after 16–24 h in culture. The C10-MJ2 cells were superinduced¹⁸ by the addition of 50 µg ml⁻¹ of cycloheximide to the culture for 4 h, 16 h after the addition of PHA and TPA. The mRNA from the induced and superinduced C10-MJ2 cells was microinjected in *Xenopus laevis* oocytes as previously described⁹. H23-cell-conditioned medium was collected 70–90 h after culturing the cells in the presence of 40 ng ml⁻¹ TPA and the BRL-conditioned medium, used at a 20-fold concentration, was as described elsewhere¹⁶. The cos-cell-conditioned medium samples were prepared as described in the Fig. 1 legend.

express high levels of growth factor activity for DA-1a cells (Table 1). (DA-1a cells that were derived from DA-1 cells were originally designated DA-2 (ref. 7), a designation we are now changing to avoid confusion with a previously described DA-2 (ref. 8) cell line.) We used mRNA from C10-MJ2 cells to isolate a cDNA clone encoding HILDA through functional expression cloning⁹ in mammalian cells (Fig. 1). This plasmid, designated pC10-6R, was a partial clone containing essentially only the protein-coding region and was used to isolate a complete 3.8 kilobase (kb) cDNA pC10-6R (ref. 4). Both of these cDNAs encode a 202 amino-acid protein which is essentially identical to the human LIF protein predicted from the human gene sequence².

RNA blot analysis showed that lectin-stimulated C10-MJ2 cells express a predominant 4 kb and a minor 1.8 kb mRNA which hybridized with the LIF/HILDA probe (Fig. 2). Both transcripts increased from very low or undetectable levels in the uninduced cells to readily detectable levels after lectin and phorbol ester stimulation. The levels of the transcripts were not significantly affected by additional treatment with cycloheximide. Similar analysis of other LIF/HILDA sources including lectin-stimulated human peripheral blood lymphocytes and phorbol ester-stimulated human lung carcinoma cells, NCI-H23 (ref. 10), confirmed that the predominant human transcript is 4 kb in size and that the minor 1.8 kb species is typically 10- to 20-fold less abundant (data not shown). Hybridization of the LIF/HILDA probe under lower stringency to blots of mRNA from one rodent cell line, BRL 32 (Buffalo rat liver; ATCC #CRL 1442), revealed the presence of several different transcripts including species that co-migrated with the 4 kb and 1.8 kb human mRNAs (Fig. 2). These results contrast with the initial report that the murine LIF transcript expressed by murine T-cell lines is only 0.75 kb (ref. 1) but are consistent with the more recent report that the predominant murine transcript is 4.5 kb (ref. 11).

The recombinant LIF/HILDA protein secreted by cos-1 cells displays the size heterogeneity typical of many glycoproteins with molecular species of relative molecular masses (M_r)

Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells

Jean-Francois Moreau*, Debra D. Donaldson, Frances Bennett, JoAnn Witek-Giannotti, Steven C. Clark & Gordon G. Wong

Genetics Institute Inc., 87 Cambridge Park Drive, Cambridge, Massachusetts 02140, USA

* Institut National de la Santé et de la Recherche Médicale (U 211-Unité de Recherche sur les Effecteurs Lymphocytaires T), Faculté de Médecine; Centre National de la Recherche Scientifique; and Centre Hospitalier Régional de Nantes, France

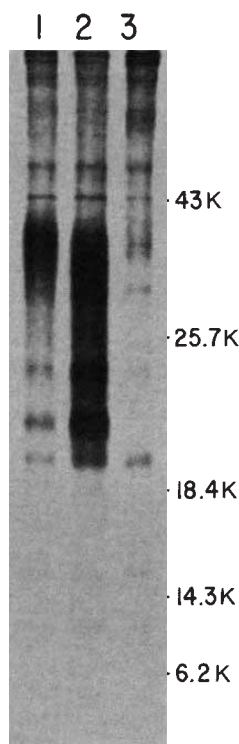
Leukaemia inhibitory factor (LIF) is a cytokine that induces macrophage differentiation of the murine M1 myeloid leukaemia cell line^{1,2}. We have isolated a cDNA clone encoding a novel human haemopoietic growth factor, human interleukin for DA cells (HILDA) that supports the proliferation of the murine interleukin-3-dependent leukaemic cell line, DA-1a (refs 3–5). HILDA proved to be identical to LIF. The demonstration that the differentiation factor LIF will also serve as a growth factor for at least one myeloid leukaemic cell line provides further evidence that the distinction between growth-promoting and differentiation-inducing activities are largely determined by the target cell type.

Analysis of the conditioned medium from a variety of human cell lines revealed that lectin-stimulated C10-MJ2 (ref. 6) T cells

The difficulty in uncoupling the processes of proliferation and differentiation in normal haemopoietic cells has led to the use of various myeloid leukaemic cell lines to analyse the relative effects of the different haemopoietic regulatory molecules separately on cell growth and cell development. These lines are believed to represent immortalized progenitor cells that are blocked at various stages of development. In the case of WEHI-3B D⁺ cells which grow in the absence of exogenous growth factors, both granulocyte colony-stimulating factor (G-CSF)¹² and granulocyte/macrophage (GM)-CSF¹³ induce the leukaemic cells to differentiate. The factor-independent M1 cells differentiate in response to G-CSF¹⁴, interleukin-6 (IL-6)¹⁵ and LIF/HILDA¹, but not GM-CSF or IL-3. The factor-dependent DA-1a cells grow without differentiating in response to IL-3, GM-CSF, G-CSF⁵ and, as shown here, LIF/HILDA, a molecule previously described as a specific differentiation factor of myeloid leukaemic cells¹. In each case, the individual myeloid cell regulatory molecules can serve either as growth-promoting

Fig. 3 SDS-polyacrylamide gel analysis of LIF/HILDA protein expressed by transfected cos-1 cells. Samples of conditioned medium from cos-1 cells pulse-labelled with [35 S]cysteine following transfection with the 3.8 kb LIF/HILDA cDNA (pC10-6R; lane 1), the 0.7 kb LIF/HILDA cDNA (pC10-6R(2), lane 2), and with pXM vector alone (lane 3) were analysed by SDS-PAGE followed by autoradiography.

Methods. The cos-1 cells were transfected and pulse-labelled as described elsewhere⁹ except that 0.5 mCi of [35 S]cysteine was used instead of methionine as label. 10 μ l of each sample were analysed under reducing conditions using a 12% polyacrylamide gel with fluorographic enhancement as previously described. The relative sizes were estimated from the mobilities of standard proteins (low relative molecular mass standards, Pharmacia).



or differentiation-inducing signals depending on the target cells. In an accompanying paper¹⁶, data are presented that show that LIF/HILDA allows the proliferation of murine embryonic stem cells while maintaining their totipotency by inhibiting their differentiation in culture. Together, these studies show that mechanisms regulating cellular differentiation and proliferation are inter-related; single regulatory molecules such as LIF/HILDA exert profoundly different effects on the two processes with different target cells. The various leukaemic and embryonal carcinoma cell lines will provide useful systems for the analysis of the interactions of LIF/HILDA with its receptor and how this interaction triggers different responses in different cell types.

The authors gratefully acknowledge R. Kriz, A. Ciarletta, N. Freddi and J. Calvetti for expert technical assistance, M. Stahl for providing the BRL-conditioned medium and F. Evers for assistance in preparing the manuscript.

Received 12 September; accepted 11 November 1988.

- Gearing, D. P. *et al.* *EMBO J.* **6**, 3995-4002 (1987).
- Gough, N. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 2623-2627 (1988).
- Moreau, J. F., Bonneville, M., Petrat, M. A., Jacques, Y. & Soullilo, J. P. *Annu. Rev. Pathol. Immun.* **137**, 25-37 (1986).
- Moreau, J.-F. *et al.* *J. Immun.* **138**, 3844-3849 (1987).
- Godard, A. *et al.* *Blood* **71**, 1618-1623 (1988).
- Arya, S. K., Wong-Staal, F. & Gallo, R. C. *Science* **223**, 1086-1087 (1984).
- Moreau, J. F. *et al.* *J. clin. Invest.* **78**, 874 (1986).
- Ihle, J. N., Rein, A. & Mural, R. in *Advances in Viral Oncology* Vol. 4 (ed. Klein, G.) (Raven, New York, 1984).
- Yang, Y. C. *et al.* *Cell* **47**, 3-10 (1986).
- Gazdar, A. F. *et al.* *Cancer Res.* **40**, 3502-3507 (1980).
- Gearing, D. P., King, J. A. & Gough, N. M. *Nucleic Acids Res.* **16**, 9857 (1988).
- Nicola, N. A., Metcalf, D., Masumoto, M. & Johnson, G. R. *J. biol. Chem.* **258**, 9017-9023 (1983).
- Metcalf, D. *Int. J. Cancer* **25**, 616-623 (1979).
- Tomida, M., Yamamoto-Yamaguchi, Y., Hozumi, M., Okabe, T. & Takaka, F. *FEBS Lett.* **207**, 271-275 (1986).
- Shabo, Y. *et al.* *Blood* (in the press).
- Smith, A. G. *et al.* *Nature* **33**, 688-690 (1988).
- Wong, G. G. *et al.* *J. Immun.* **140**, 3040-3044 (1988).
- Shaw, J., Meerovitch, K., Elliott, J. F., Blackley, R. C. & Paetkau, V. *Molec. Immun.* **24**, 409-419 (1987).
- Shaw, G. & Kamen, R. *Cell* **46**, 659-667 (1986).
- Messing, J. & Vieira, J. *Gene* **19**, 269-276 (1982).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Direct interaction between fos and jun nuclear oncoproteins: role of the 'leucine zipper' domain

Paolo Sassone-Corsi, Lynn J. Ransone, William W. Lamph & Inder M. Verma

Molecular Biology and Virology Laboratory, The Salk Institute, PO Box 85800, San Diego, California 92138, USA

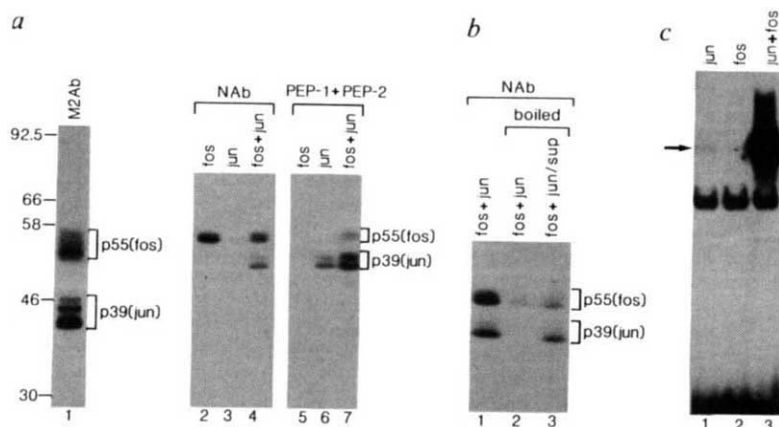
Gene expression is modulated by the specific interactions of nuclear proteins with unique regulatory sequences in the genome. Proteins involved in transcriptional regulation seem to be either transcription factors or transcription modulators and their interactions are crucial in determining whether the expression of a specific gene is activated or repressed. Recently¹, the product of the proto-oncogene *jun* has been identified as the transcription factor AP-1, whereas nuclear oncoproteins *fos* and *myc* have been implicated in transcriptional transregulation of several promoters²⁻⁶. Furthermore, the products of the *fos* and *jun* proto-oncogenes are associated in some transcription complexes^{3,5,7}. Although the nature of the association is unclear, the two proteins co-immunoprecipitate with *fos* antibodies in nuclear extracts⁸⁻¹⁰. Here, we report studies that demonstrate that the *fos* protein directly modulates *jun* function by means of a heterodimer of *fos* and *jun* proteins. The *fos* 'leucine zipper'¹¹ domain is necessary for the DNA binding of the heterodimer; a distinct domain, localized in the C-terminal region of the *fos* protein, is responsible for transcriptional regulation.

As the *fos* protein seems to have DNA-binding properties only in the presence of a cellular extract^{12,13}, its pleiotropic regulatory characteristics are probably due to association with other transcription factors. One of the *fos*-associated polypeptides, the nuclear protein p39 is, in fact, the product of the *jun* proto-oncogene, the transcription factor AP-1⁸⁻¹⁰. Interestingly, the *fos* gene product enhances *c-jun* activation function of a TPA (12-*O*-tetradecanoyl-phorbol-13-acetate) responsive promoter element (TRE), providing evidence in favour of the hypothesis of interaction and cooperativity between the two nuclear oncoproteins¹⁰. To test whether other nuclear components were required for the association, we synthesized *jun* and *fos* proteins by *in vitro* translation using RNA transcribed *in vitro* from the bacteriophage T7 promoter linked to either the *fos* or *jun* full-length cDNAs^{14,15}. Both proteins are of the expected apparent relative molecular mass and seem to be modified in the same way as the *in vivo* products (Fig. 1a, compare lane 1 with lanes 2 and 6). When the two proteins were co-translated or mixed after translation and then immunoprecipitated with *fos* antibodies, they seem to associate with the same molar ratio as *in vivo* (Fig. 1a, compare lane 1 with lane 4). The reciprocal experiment is not quantitative because *jun* antibodies do not immunoprecipitate all the *fos*-*jun* complex (compare lanes 4 and 7; ref. 10). In control experiments, reticulocyte lysate alone had no effect on the association of the *in vitro*-synthesized *fos* or *jun* proteins (data not shown). The *in vitro* association of *fos* and *jun* can be abolished by boiling the mixed sample before immunoprecipitation (see Fig. 1b). This result is analogous to that obtained using *in vivo*-synthesized proteins^{9,10}.

To test whether the *jun* protein synthesized *in vitro* can bind to a TRE, we used DNA-protein binding studies and analysed the nucleoprotein complex on non-denaturing retardation gels. As shown in Fig. 1c, *jun* synthesized *in vitro* binds with a very low affinity to a TRE oligonucleotide (Fig. 1c; lane 1). The affinity of binding to the TRE is dramatically enhanced by the addition of *in vitro*-synthesized *fos* protein (Fig. 1c; lane 3). The *fos* protein alone has no binding activity (Fig. 1c; lane 2). Titration experiments established that the increase in

Fig. 1 Association of *in vitro*-synthesized fos and jun proteins. **a**, Lane 1, an immunoprecipitation of the p55^{fos}-p39^{jun} complex from serum-stimulated NIH3T3 cells using the fos M2 antibody¹⁶. Fos and jun proteins were synthesized separately *in vitro* (lanes 2, 3, 5, 6) or together (lanes 4 and 7) and immunoprecipitated with fos monoclonal antibodies (NAb) (lanes 2-4) or PEP-1 and PEP-2 jun antibodies¹ (lanes 5-7). Relative molecular masses are on the left, in thousands. **b**, Fos and jun proteins were synthesized *in vitro*, mixed and immunoprecipitated with NAb¹⁵ (lane 1); the complex was boiled before immunoprecipitation showing dissociation of fos from jun in the immunoprecipitate (lane 2), whereas the supernatant (sup) still contains both proteins (lane 3). **c**, DNA-binding activity of the *in vitro*-synthesized proteins analysed by gel-shift assay using an oligonucleotide containing the human metallothionein IIA TRE¹⁷. The arrow indicates the specific retarded bands as shown by competition experiments using excess of unlabelled TRE fragment (data not shown); the lower band is related to non-specific binding of proteins from the reticulocyte lysate. In lanes 1 and 2 the equivalent of 8 μ l of jun- or fos-containing lysate were used for the binding reactions; in lane 3, 1 μ l of each was pre-mixed and then associated with the labelled TRE fragment.

Methods. For the immunoprecipitation of the fos-jun complex, serum-stimulated NIH3T3 cells were labelled with 0.5–1 mCi ml⁻¹ of *trans*-³⁵S-label (ICN Radiochemicals) for 90 min. Proteins from nuclear lysates were immunoprecipitated by the addition of 2 μ l of anti-fos M2 antibody and analysed by 10% SDS-PAGE as described¹⁸. *jun* and *fos* RNAs were synthesized *in vitro* as described after insertion of the respective full cDNA sequences under the control of the T7 bacteriophage promoter¹⁹. About 2 μ g of *fos* and *jun* RNAs were mixed with 70 μ l of reticulocyte lysate (Promega) to obtain proteins synthesized *in vitro*, in a final volume of 100 μ l. When fos and jun proteins were synthesized in the same translation reaction about 2 μ g of each RNA was used. About 1/10 of the *in vitro*-translated proteins were used for each immunoprecipitation. The gel-shift analysis was performed as described¹⁰. Various amounts of the *in vitro* translated proteins (typically 1–5 μ l) were incubated with 2 μ g of poly (dI:dC) (Boehringer) in TM buffer (50 mM Tris-HCl, pH 7.9, 12.5 mM MgCl₂, 1 mM EDTA, 1 mM dithiothreitol, 20% glycerol) for 20 min at room temperature in a 20- μ l final volume. A synthetic 18-base pair oligonucleotide containing the human metallothionein IIA TRE was end-labelled with [³²P]ATP, using T4 polynucleotide kinase. Approximately 0.1 ng of ³²P-labelled DNA (<10,000 c.p.m.) was added to the preincubated proteins. DNA-protein complexes were resolved on a 4% polyacrylamide gel (39:1 acrylamide to bisacrylamide) in 0.25 $\times$ TBE (1 $\times$ TBE: 50 mM Tris-Borate, pH 8.3, 1 mM EDTA). The gels were dried and autoradiographed with intensifying screens at -70 $^{\circ}$ C.



TRE-binding activity observed when the fos protein was added is of the order of 1,000-fold (data not shown). Thus, the formation of a heterocomplex between fos and jun is required for increased DNA-binding activity. We believe that the fos and jun proteins associate to form a heterodimer because the two proteins associate in a 1:1 molar ratio (Fig. 1) and because on sucrose gradients the fos-jun complex sediments faster than the fos monomer (data not shown).

To investigate further the interactions between the two proteins, we used fos and jun antibodies to inhibit complex formation on the TRE. Figure 2b shows that addition of polyclonal fos antibody (M2Ab; lane 5) and jun antibody PEP-1 (lane 8) strongly inhibited binding of the fos-jun protein complex to the TRE; the monoclonal anti-fos antibody (NAb) and jun antibody PEP-2 had a less pronounced effect on DNA binding (lanes 6 and 9). The strong inhibition of DNA-binding by M2Ab (lane 5) is interesting because the recognition site of M2Ab is near the fos 'leucine zipper' domain (Fig. 2a). Similarly, the effect of PEP-2 antibody was less pronounced because the recognition site is removed from the presumed DNA-binding domain. Even though the monoclonal antibody did not prevent the fos-jun complex binding DNA, the complex did show a decreased mobility, suggesting that the N-terminal end of the fos protein (amino acids 4–17) may be free to bind to the NAb and thus increase the apparent size of the fos-jun complex resulting in a reduced mobility (lane 6).

Combinations of anti-fos and anti-jun antibodies were used to investigate the structural characteristics of the fos-jun heterodimer binding to the TRE. As shown in Fig. 2c, addition of either PEP-1 and PEP-2 further decreased the binding already impaired by the fos M2Ab (lanes 2–4). Because PEP-1 alone has a stronger effect on binding than PEP-2 (see Fig. 2b, lanes 8 and 9), the combination M2Ab/PEP-1 has more effect on the formation of the complex than M2Ab/PEP-2 (Fig. 2c; lanes 3 and 4). Similarly, the combination NAb/PEP-1 inhibits binding more than NAb/PEP-2 (Fig. 2c; lanes 6 and 7).

We conclude that the domains of fos and jun proteins that are implicated in DNA binding are near the 'leucine zipper' region.

Because of the hypothetical role of 'leucine zipper' in dimerization¹¹, we constructed two truncated fos proteins, one containing the 'leucine zipper' (fos/198) and the other in which this domain is removed (fos/160) (Fig. 3a). The two truncated proteins (fos/160 and fos/198) were synthesized *in vitro* and visualized by SDS-PAGE. They show the expected reduction in relative molecular mass (Fig. 3b). The truncated fos proteins in association with jun were tested for DNA binding to the TRE (Fig. 3c). The results indicate that fos/160 does not cooperate with jun, as binding is completely abolished (lanes 8–10), whereas the fos/198 truncated protein cooperates with jun as efficiently as the wild-type fos protein (compare lanes 5–7 with lanes 11–13, also notice the increased mobility of fos/198 complex). The experiments in Fig. 3c demonstrate that the fos 'leucine zipper' domain is essential for cooperativity with jun in binding to a TRE. Furthermore, all fos protein sequences downstream from position 198 are dispensable for DNA-binding activity. We therefore wanted to determine whether additional fos protein domains would be required for the cooperativity with jun in transcriptional activation. We have already shown that an increase in transcription is observed when TRE/thymidine kinase-chloramphenicol acetyl transferase (TK-CAT) reporter plasmid is co-transfected with a c-jun expression vector in F9 embryonal carcinoma cells¹⁰. This activation, due to the addition of exogenous jun protein, is enhanced by the co-transfection of a fos expression vector^{9,10}. We therefore tested the truncated forms of fos in the same *in vivo* co-transfection assay. When both fos and jun expression vectors were co-transfected, the expected increase was observed (lane 6). The expression of either fos/198 or fos/160 showed no transcriptional induction (lanes 3 and 4); the two truncated fos proteins also failed to cooperate with c-jun (lanes 7 and 8). Both the truncated fos/160 and fos/198 were synthesized as judged by

Fig. 2 Effect of fos and jun antibodies on the DNA binding of the fos-jun complex. *a*, Schematic representation of fos and jun proteins deduced from the nucleotide sequence and indicating the major features of both proteins. Vertical stripes indicate the serine sites of phosphoesterifications in the fos protein¹⁸; hatched box, leucine-zipper; solid box, region of homology to GCN4. The positions of the peptides that anti-fos N monoclonal¹⁵ and M2 polyclonal¹⁶ antibodies were generated against, and the position of PEP-1 and PEP-2 jun peptides are shown¹. aa, amino acid. *b*, Jun and fos *in vitro*-translated proteins are tested for DNA binding to a TRE. The effect of either the fos or the jun antibody is tested on the binding of the fos-jun complex. *c*, Analysis of different combinations of various antibodies on the binding of the fos-jun complex. Lane 1 is equivalent to lane 2 in *b*. **Methods.** Binding reactions and gel retardation were performed as described in the Fig. 1 legend. Antibody treatment was as described^{5,10}. The immunoreactions were typically carried out at 0 °C for 1–2 h. In these experiments both fos and/or jun antibodies were added to the pre-formed jun-fos complex, before the DNA binding reaction. Mixing the antibodies to the respective antigen before the association reaction inhibits the DNA binding activity slightly more (data not shown).

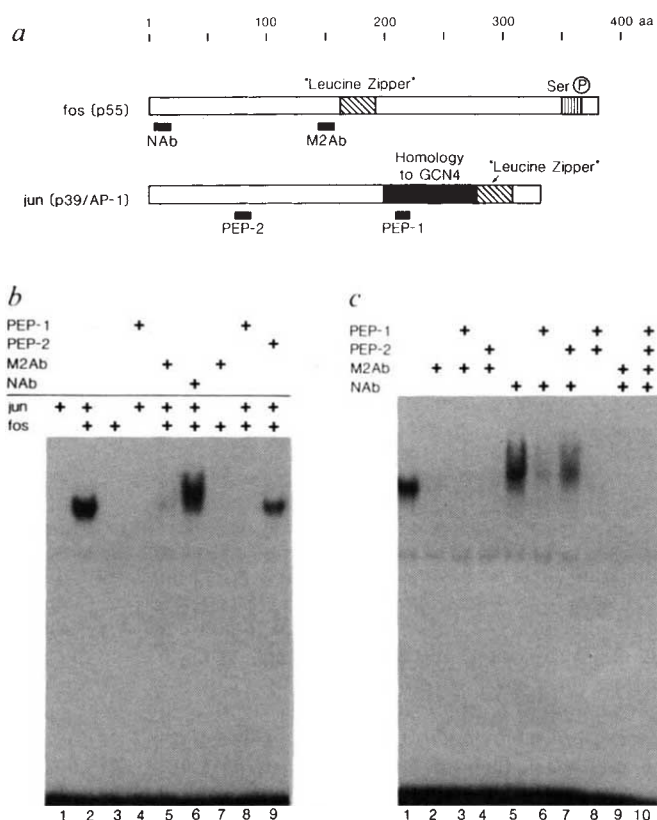
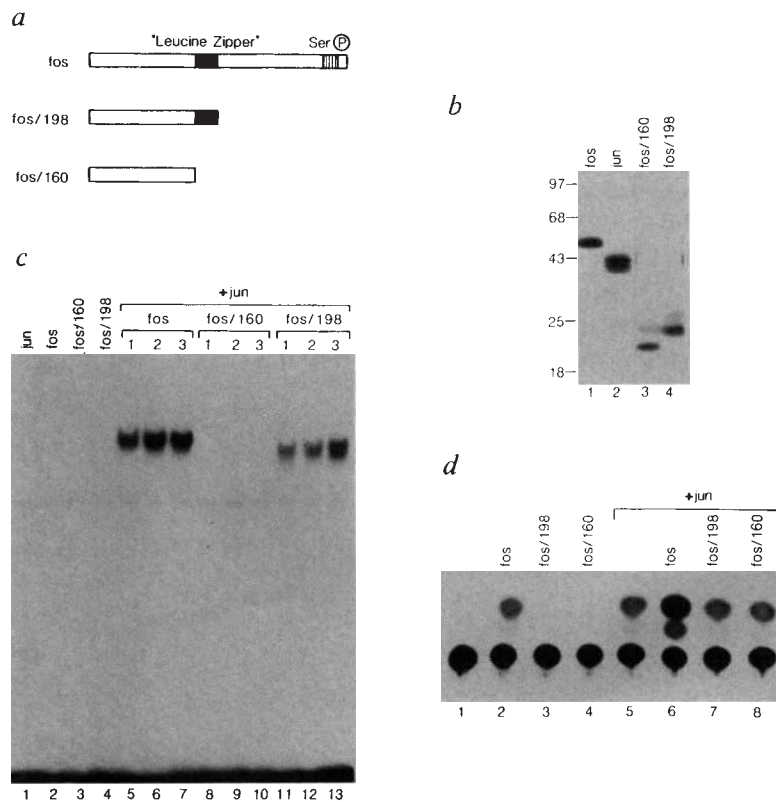


Fig. 3 The fos 'leucine zipper' domain is required for DNA binding but is not sufficient for transactivation function. *a*, Truncated forms of the fos protein were obtained using *Bam*HI linker-insertion mutants in *Alu*I sites of the human *c-fos* cDNA at amino-acid positions 160 and 198 (ref. 20). The fos 'leucine zipper' domain (solid box) is localized between amino acids 165 and 195 (ref. 11). Vertical stripes indicate serine sites as in Fig. 2. *b*, *In vitro* synthesis of the wild-type fos and jun proteins and of the mutants fos/198 and fos/160. Relative molecular masses are on the left, in thousands. *c*, DNA-binding assay using a TRE oligonucleotide and the truncated fos proteins. Lanes 4–6 show binding activity of the heterocomplex fos/jun; in this experiment the amount of *in vitro*-synthesized jun is constant (1 μ l), whereas fos protein is increasing (1–3 μ l). *d*, Undifferentiated F9 embryonal carcinoma cells were co-transfected with reporter plasmid TRE/TK-CAT and a combination of *fos* and *jun* plasmids. Recombinant TRE/TK-CAT contains a human metallothionein IIA TRE sequence (the TRE consensus is GTGACTCAG) inserted upstream of the herpes simplex TK promoter (positions –109 to +57) which is fused to the CAT gene¹⁷. Full-length *c-fos* expression vector (BK28) has been previously described¹⁰.

Methods. Transfection of plasmid DNA into F9 embryonal carcinoma cells used the calcium phosphate co-precipitation technique. Cells were passaged in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. Cells were transfected when 70% confluent and exposed to the precipitate for 12–16 h. Equal molar ratios (1:1 in lanes 1–4 and 1:1:1 in lanes 5–8) of reporter plasmid and trans-activator vector were transfected, to obtain total of 20 μ g of DNA per 10-cm culture dish.



nuclear staining with the monoclonal anti-fos antibody (data not shown).

Our results indicate that the fos 'leucine zipper' domain (amino acids 165–193) seems to be essential for binding activity. However, other protein regions must be important in conferring the high-affinity binding to the target sequence. For instance fos

protein alone does not bind to a TRE, and jun protein binds very poorly despite the intact DNA-binding domain and 'leucine zipper' (Fig. 1c). The data presented here also supports our previous notion that the C-terminal domain of the fos protein, which is extensively modified by serine phosphoesterifications, is required for transcriptional transregulation⁵.

The formation of a presumptive fos-jun heterodimer probably dictates a non-symmetrical conformation of the complex. Homodimers are rotationally symmetric molecules, whether they associate in a parallel or anti-parallel way. There is little physical evidence about mode of association of the fos-jun heterodimer, but we anticipate that it might be crucial to the pleiotropism of fos function. The association of two oncoproteins, fos and jun, constitutes a model system which could serve as a paradigm for studying the molecular links between signal transduction and transcriptional regulation. Further insights would require extensive mutational analysis of both fos and jun proteins.

We are grateful to T. Bos and P. Vogt for the generous gift of the anti-jun PEP-1 and PEP-2 antibodies. We thank P. Wamsley for technical assistance and P. McClintock for help in preparation of the manuscript. P. S.-C. is supported by the American Cancer Society and is on leave from the Laboratoire de Genetique Molculaire des Eucaryotes, CNRS, Strasbourg, France. L.J.R. and W.W.L. are supported by the NIH; I.M.V. by the NIH and the American Cancer Society.

Received 5 October; accepted 24 October 1988.

- Bohmann, D. *et al. Science* **238**, 1386-1392 (1987).
- Setoyama, C., Frunzio, R., Liao, G., Mudryj, M. & DeCrombrughe, B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3213-3217 (1986).
- Distel, R. J., Ro, H.-S., Rosen, B. S., Groves, D. L. & Spiegelman, B. M. *Cell* **49**, 835-844 (1987).
- Rauscher III, F. J., Sambucetti, L. C., Curran, T., Distel, R. J. & Spiegelman, B. M. *Cell* **52**, 471-480 (1988).
- Sassone-Corsi, P., Sisson, J. C. & Verma, I. M. *Nature* **334**, 314-319 (1988).
- Kingston, R. E., Baldwin, A. S. Jr & Sharp, P. A. *Nature* **312**, 280-282 (1984).
- Franza, B. R. Jr, Rauscher III, F. J., Josephs, S. F. & Curran, T. *Science* **239**, 1150-1153 (1988).
- Rauscher III, F. J. *et al. Science* **240**, 1010-1016 (1988).
- Chiu, R. *et al. Cell* **54**, 541-552 (1988).
- Sassone-Corsi, P., Lamph, W. W., Kamps, M. & Verma, I. M. *Cell* **54**, 553-560 (1988).
- Landschulz, W. H., Johnson, P. F. & McKnight, S. L. *Science* **240**, 1759-1764 (1988).
- Sambucetti, L. C. & Curran, T. *Science* **234**, 1417-1419 (1986).
- Renz, M., Neuberger, M., Kurz, C., Bravo, R. & Muller, R. *EMBO J* **4**, 3711-3716 (1985).
- Lamph, W. W., Wamsley, P., Sassone-Corsi, P. & Verma, I. M. *Nature* **334**, 629-631 (1988).
- De Togni, P., Niman, H., Raymond, V., Sawchenko, P. & Verma, I. M. *Molec. cell. Biol.* **8**, 2251-2256 (1988).
- Curran, T., Van Beveren, C., Ling, N. & Verma, I. M. *Molec. cell. Biol.* **5**, 162-172 (1985).
- Angel, P. *et al. Cell* **49**, 729-739 (1988).
- Barber, J. R. & Verma, I. M. *Molec. cell. Biol.* **7**, 2201-2211 (1987).
- Melton, D. *et al. Nucleic Acids Res.* **12**, 7035-7056 (1984).
- Sassone-Corsi, P., Lamph, W. W. & Verma, I. M. *Cold Spring Harb. Symp. quant. Biol.* **53**, (in the press).

Receptors bound to antiprogesterin form abortive complexes with hormone responsive elements

Anne Guiochon-Mantel*, Hugues Loosfelt*,
Thierry Ragot†, Alain Bailly*, Michel Atger*,
Micheline Misrahi*, Michel Perricaudet†
& Edwin Milgrom*‡

* Hormones et Reproduction (INSERM U135), Faculté de Médecine, Paris-Sud, 94275 Le Kremlin Bicêtre Cedex, France
† Laboratoire de Génétique des Virus Oncogènes (CNRS UA 220), Institut Gustave Roussy, PR11, 94805 Villejuif Cedex, France

The mechanism of action of antisteroids is not understood and explanations of their antagonistic activity have been sought at all levels of hormone action. It has been proposed that antisteroids, after binding to receptor, trap it into a non-activated (non DNA-binding) form possibly through interaction with a heat-shock protein of relative molecular mass (M_r) 90,000 (90 K)¹, or that the antisteroids provoke binding of receptor to nonspecific DNA sites but not to hormone responsive elements (HREs)², or that the antisteroid-receptor complexes can bind to HREs but form abortive

complexes that fail to regulate transcription^{3,4}. We have constructed a deleted cDNA encoding a mutant form of rabbit progesterone receptor which exhibits constitutive activity, that is, binds to HREs in the absence of hormone and thus bypasses the first two steps discussed above. Co-transfection experiments allowed the expression of both constitutive and wild-type receptors in the same recipient cells. Antiprogesterin RU486-wild-type receptor complexes completely suppressed the activity of the constitutive receptor on a reporter gene, showing that the inhibition is at the level of their common responsive elements.

Wild-type or deleted forms of rabbit progesterone receptor cDNA were inserted into the expression vector pKSV10 (Pharmacia) and used to transfect COS-7 cells⁵. The biological activity of the receptors was tested on a reporter gene containing the HRE of mouse mammary tumour virus upstream from the chloramphenicol acetyltransferase structural gene (MMTV-CAT)⁶. A total deletion of the steroid-binding domain (amino acids 663-930) produced a constitutive receptor (60% of the activity of the wild-type receptor saturated with hormone) which was insensitive to the action of the progestin R5020 (17,21-dimethyl-19-norpregna-4,9-dien-3,20-dione) or to that of the antagonist RU486 (17 β -hydroxy-11 β -(4-dimethylamino-phenyl)-17 α -(1-propynyl)-oestra-4,9-dien-3-one), whereas the latter exerted, as expected, a strong inhibitory effect in the presence of the wild-type receptor (Fig. 1a).

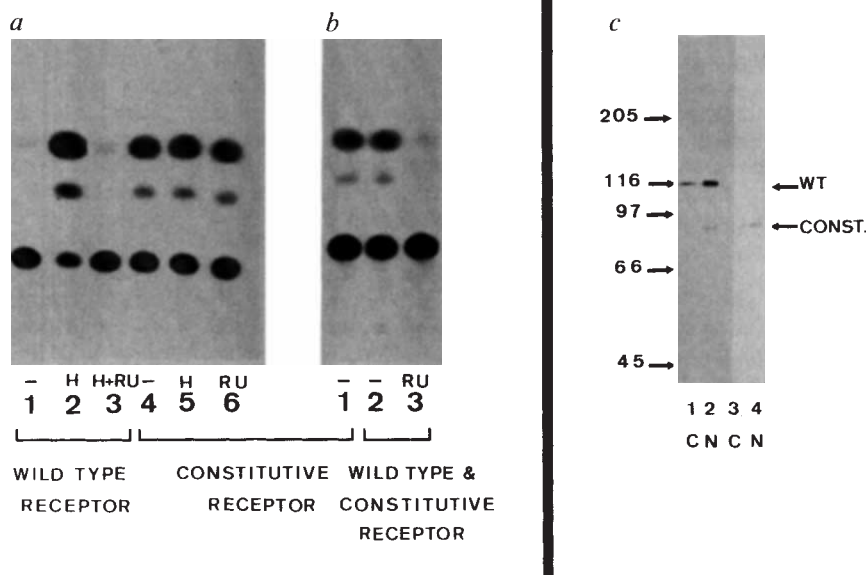
These observations suggested the possibility of testing various hypotheses on the mechanism of action of antisteroids. If the spontaneously activated constitutive receptor and the wild-type receptor-antagonist complexes do bind to the same DNA sites then the latter would inhibit the activity of the former. On the other hand, if antagonists inhibit activation or lead to binding of receptor to nonspecific DNA, they would do so on the wild-type receptor but, not binding to it, would not modify the activity of the constitutive receptor. The results of the co-transfection experiment clearly supported the first hypothesis: RU486-wild-type receptor complexes suppressed the activity of the constitutive receptor (Fig. 1b). Two types of controls were performed. RNase mapping showed that the changes in CAT activity were related to actual changes in the level of transcription originating at the proper site (data not shown). The possibility that the co-transfection with wild-type receptor cDNA and RU486 treatment had decreased, by an unknown mechanism, the concentration of constitutive receptor or its distribution between soluble and tight chromatin-binding forms was also tested. These possibilities were borne out by an immunoblot experiment in which the constitutive receptor was analysed in the cytosol and nuclear extract of cells either transfected only with its coding cDNA or also with the cDNA for wild-type receptor and treated with RU486 (Fig. 1c).

The regions of the receptor molecule which were involved in the antagonistic effect of receptor-RU486 complexes were then analysed using expression vectors encoding mutated receptors. As expected, receptors deleted in the DNA-binding (Fig. 2, pKSV-rPR 5 and pKSV-rPR 6) or in the steroid binding (Fig. 2, pKSV-rPR 7, pKSV-rPR 8 and pKSV-rPR 9) regions were devoid of inhibitory activity. In the N-terminal region the deletion of amino acids 25-103 (pKSV-rPR 3) and 373 to 546 (pKSV-rPR 4) had no effect on the inhibitory activity of receptor-antagonist complexes. Deletion of amino acids 103-546 (pKSV-rPR 2) produced a receptor form that had a low inhibitory activity—the inhibition was more apparent when competing mutant receptor concentration was increased fivefold (data not shown).

The possibility existed, however, that the inhibitory activity of receptor-RU486 complexes was a low-affinity, nonspecific, phenomenon only observed at very high concentrations of these complexes. To test this, and to obtain a rough estimate of the stoichiometry of the receptor molecules involved in the inhibition reaction, two types of experiments were performed. The ratio of constitutive/wild-type receptor was varied by changing

‡ To whom correspondence should be addressed

Fig. 1 *a*, The antiprogesterin RU486 abolishes induction of MMTV-CAT reporter gene by R5020 in the presence of wild-type receptor and has no effect on the constitutive induction of MMTV-CAT by mutated receptor. COS.7 cells were transfected either with pKSV-rPR 0 encoding wild-type receptor (lanes 1–3), or with pKSV-rPR 1, encoding constitutive receptor (lanes 4–6) and were treated with vehicle alone (–), 10 nM R5020 (H), 1 μ M RU486 (RU) or 10 nM R5020 + 1 μ M RU486 (H+RU). *b*, Constitutive activation of MMTV-CAT by pKSV-rPR 1 is abolished by pKSV-rPR 0 in the presence of RU486. COS.7 cells were transfected either with pKSV-rPR 1, (lane 1) or with pKSV-rPR 0 and pKSV-rPR 1 (lanes 2 and 3) and were treated with vehicle alone (–) or 1 μ M RU486 (RU). Similar results were observed with CV1 cells (not shown). *c*, The concentration and subcellular distribution of constitutive receptor is not modified by the presence of wild-type receptor and RU486 treatment. COS.7 cells were either transfected with 1 μ g of pKSV-rPR 1, encoding the constitutive receptor (lanes 3, 4) or transfected with 1 μ g of pKSV-rPR 1 and 10 μ g of pKSV-rPR 0 and treated with 1 μ M RU486 (lanes 1, 2). Cells were collected and cytosolic (C) and nuclear (N) extracts were prepared and analysed by immunoblot. Migrating wild-type receptor (WT) and constitutive receptor (CONST.) is shown on the right and relative molecular masses are shown on the left, in thousands.



Methods. The entire open reading frame of the cDNA encoding the rabbit progesterone receptor¹⁹, from –3 to 2,900, was inserted at the *Bgl*II site of pKSV10 (Pharmacia). The expression vector encoding constitutive receptor was obtained by inserting a stop codon after codon 662. The insertion site was verified by sequencing. *a*, Transfections were performed as follows. COS.7 cells were grown in Dulbecco's Modified Eagle Medium supplemented with 10% charcoal-treated FCS two days before transfection. 1×10^6 – 2×10^6 cells were transfected using the calcium phosphate precipitate method²⁰ with 10 μ g of expression vector and 10 μ g of reporter plasmid MMTV-CAT. After 18 h, the precipitate was removed and replaced by fresh medium supplemented with 10% charcoal-treated FCS containing the hormone as indicated. Cells were collected after 24 h. Cytosolic extracts were prepared and CAT assays were run for 45 min using 100 μ g of protein as described²¹. The same results were obtained using various concentrations of expression vector DNA (0.5–10 μ g). *b*, Similar transfection experiments were carried out except that 1 μ g of each expression vector was used and carrier DNA was added to adjust the total DNA to 20 μ g per plate. *c*, Transfected cells were homogenized in 3 ml g^{-1} . Cytosol, and nuclear extracts (in 0.4 M NaCl buffer) were prepared as described^{22,23}. They were precipitated with four volumes of acetone and loaded on a 9% SDS-polyacrylamide gel. Proteins were transferred to nitrocellulose membrane and probed with Mi60-10 anti-receptor monoclonal antibody²⁴, followed by incubation with anti-mouse immunoglobulin antibodies and ¹²⁵I-labelled protein A as described elsewhere²³.

Fig. 2 Inhibition of the activity of the constitutive receptor by different mutated forms of receptor in the presence of RU486. *a*, Receptor deletion mutants and their structure relative to the wild-type protein. The DNA-binding region is represented by the stippled area. In front of each mutant the percentage activity relative to MMTV-CAT reporter gene without (–H) and with (+H) 10 nM R5020 is shown. *b*, Inhibitory activity. COS.7 cells were co-transfected with 1 μ g of the constitutive receptor expression vector (pKSV-rPR 1) and 1 μ g of the competitor (wild-type or mutated receptor) along with 10 μ g of MMTV-CAT and 5 μ g of control plasmid pCH110 (Pharmacia), and treated either with vehicle alone (–) or with 1 μ M RU486 (+) 24 h before collection.

Methods. Each mutant was constructed by cleavage at the proper restriction sites, ligation and reinsertion in the expression vector pKSV-rPR. All the constructions were characterized by sequencing of these junctions. The following amino acids were deleted in the mutated receptors: 663–930 in pKSV-rPR 1; 103–546 in pKSV-rPR 2; 25–103 in pKSV-rPR3; 373–546 in pKSV-rPR 4; 547–662 in pKSV-rPR 5; 593–640 in pKSV-rPR 6; 662–736 in pKSV-rPR7; 720–782 in pKSV-rPR 8; and 738–930 in pKSV-rPR 9. Cells were transfected and CAT assays were performed on the cytosolic extracts as described in the Fig. 1b legend. The CAT activity was calculated for each mutant as the mean value obtained in two to four independent experiments. The percentage of chloramphenicol acetylation was determined by counting the silicagel. The β -galactosidase assays were performed as described²⁵. The percentage of chloramphenicol converted per unit of β -galactosidase activity was calculated for each mutant and compared to that of the expression vector encoding wild-type receptor.

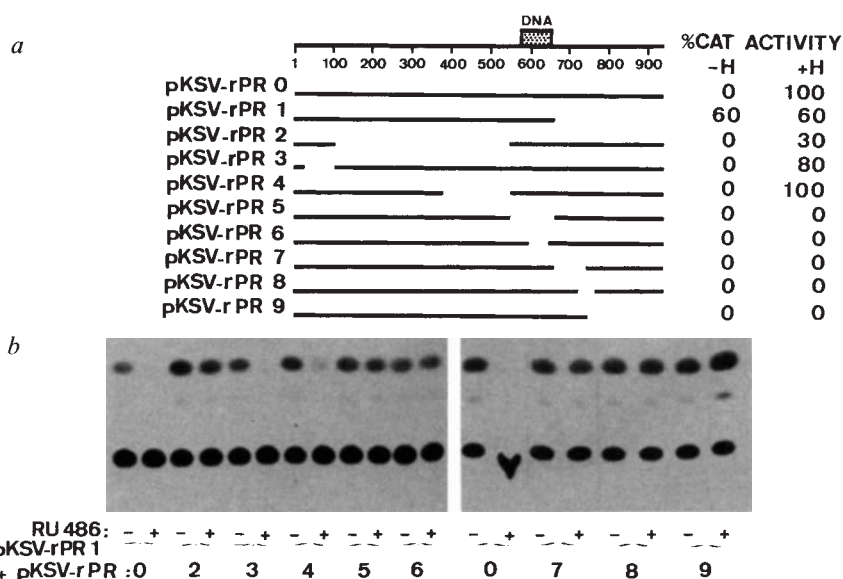
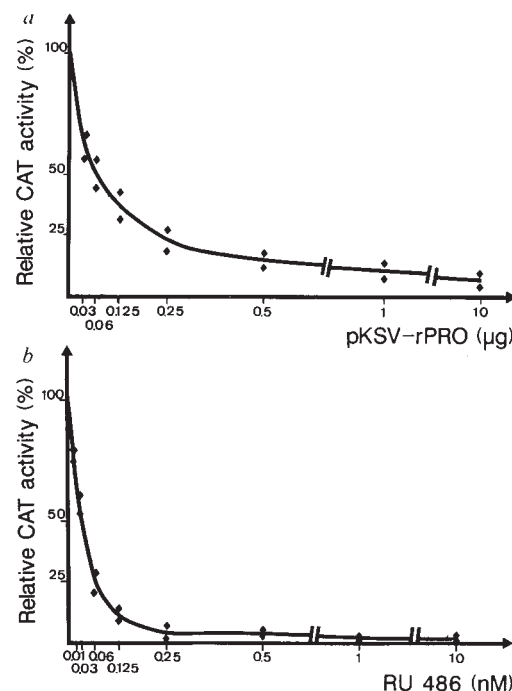


Fig. 3 Inhibition of the activity of the constitutive receptor by various concentrations of wild-type receptor (*a*) and RU486 (*b*). *a*, COS.7 cells were co-transfected with 1 μ g of pKSV-rPR 1 (constitutive) and increasing amounts (0–10 μ g) of pKSV-rPR 0 (wild-type) along with 10 μ g of MMTVCAT and 5 μ g of pHC110. Cells were treated with 1 μ M RU486 24 h before collection. *b*, Cells were co-transfected with 1 μ g of pKSV-rPR 0, 1 μ g of pKSV-rPR 1 along with 10 μ g of MMTV-CAT and 5 μ g of pHC110. Cells were treated 24 h before collection with increasing concentrations of RU486. **Methods.** Transfections and CAT assays were performed as described in the Fig. 1 legend. The CAT activity was determined as described in the Fig. 2 legend. (100% CAT activity refers to the percent of acetylation observed with 1 μ g of pKSV-rPR 1 and 1 μ M RU486 (*a*) or with 1 μ g of pKSV-rPR 1 and 1 μ g of pKSV-rPR 0 (*b*). For each concentration of pKSV-rPR 0 (*a*) or of RU486 (*b*) the percentage of acetylation is expressed relative to that reference).



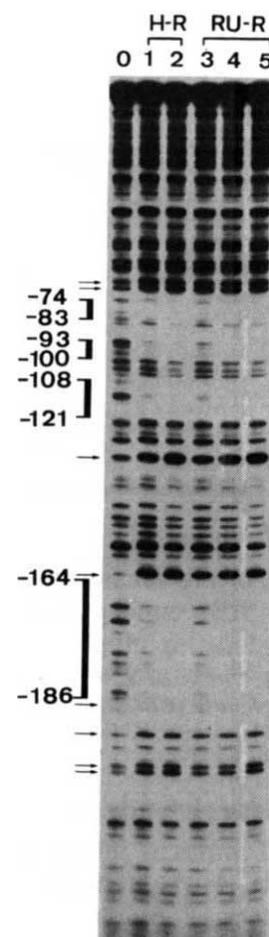
the ratio of co-transfected DNAs (Fig. 3)—radioimmunoassays of receptor protein⁷ had shown that wild-type or constitutive receptor cDNAs were expressed with similar efficiency (data not shown). Six per cent of wild-type receptor saturated with RU486 was sufficient to inhibit 50% of the activity of constitutive receptor, compared with the 60% activity of the constitutive receptor (see Fig. 1*a*). Moreover, when equal amounts of both receptor cDNAs were used for transfection, addition of 0.03 nM RU486 was sufficient to inhibit half the activity of the constitutive receptor (K_d of receptor–RU486 interaction: 0.1 nM)⁸. The low proportion of RU486–wild-type receptor complexes suppressing constitutive receptor activity may be due to the fact that the MMTV DNA present in the reporter gene contains four receptor-binding sites⁹ (each of these may interact with a receptor oligomer). It is possible that partial occupation of some of these sites by RU486–wild-type receptor complexes is sufficient to inhibit hormonal stimulation even if other sites are occupied by constitutive receptor. This would agree with results of mutation studies which showed that the four receptor binding sites were necessary for optimal hormonal activity¹⁰.

The results presented here suggest that *in vivo* antiprogesterone receptor complexes bind to HREs but fail to provoke the proper biological response. Differences in the dissociation kinetics and the sedimentation properties of agonist and antagonist–receptor complexes have been previously reported¹¹. These may however, be consequences of differing complex conformations and may not explain their opposite biological activity. Activation of antiprogesterone RU486–receptor complexes has previously been observed in cytosol¹² and we have shown that purified complexes bind to regulatory regions of the uteroglobin gene and give similar DNase footprints to those observed with agonist–receptor complexes¹³. The possibility has been raised however, that purified receptors behave in a different way from that prevailing *in vivo*¹⁴. We have repeated these experiments with HREs of MMTV and have observed a similar binding of agonist and antagonist–receptor complexes to exactly the same HREs (Fig. 4).

It is not clear if our observations on antiprogesterones may be generalized to other antisteroids. Antigluocorticoids and especially RU486 when it binds to glucocorticoid receptor, seem to provoke only a weak activation (defined as tight chromatin-or

Fig. 4 DNase I footprints showing interaction of receptor complexed to agonist (H-R) or antagonist (RU-R) with the HREs of the MMTV long terminal repeat. The non-coding strand of the 390-base pair *Hind* III–*Bgl* II fragment of the plasmid pHCwt (ref. 10) was labelled at the 5' terminus of the *Hind* III site. This fragment contains MMTV long terminal repeat sequences from –237 to +125 with respect to the transcription start point. The fragment (0.2 nM) was incubated without receptor (lane 0) or with different concentrations of hormone (H-R) or RU486 (RU-R) receptor complexes: lane 1, 26 nM; lane 2, 52 nM; lane 3, 23 nM; lane 4, 46 nM; lane 5, 69 nM. DNA was then digested with DNase I and analysed on a 6% sequencing gel. Regions protected (vertical lines) or showing enhanced DNase I cleavage (arrows) upon receptor binding are indicated. The locations of the footprints along the MMTV DNA were determined using Maxam and Gilbert sequencing reactions²⁶ (data not shown).

Methods. Purification of R5020 or RU486–receptor complexes from rabbit uterus by immunoaffinity chromatography²⁷ and DNase I footprinting experiments with purified receptor¹³ were performed as described.



DNA-binding)^{15,16}. *In vivo* footprinting experiments have shown no binding of antigluocorticoid–receptor complexes to HREs in the tyrosine aminotransferase gene¹⁷. On the contrary, the antioestrogen hydroxytamoxifen determines activation of the oestrogen receptor¹⁸. Binding to HREs of antioestrogen–receptor complexes has not been studied. It is thus possible that antisteroids fall in two categories: some impeding receptor acti-

vation, others and especially antiprogestins forming abortive complexes with receptor and HREs. The method described here may now be used to rapidly assign new compounds obtained by synthesis to each of these categories. The exact mechanisms of receptor interaction with HREs and subsequent changes in the rate of transcription are not understood. They may involve receptor dimerization or polymerization, interaction with transcription factors or RNA polymerase itself. Any one of these later events may be prevented from occurring when the receptor has bound an antagonist instead of an agonist. Analysis of transcription complexes formed in the presence of agonist and antagonist-receptor complexes should give insight into the mechanisms of hormone action following receptor binding to HREs.

We thank D. Philibert (Roussel Uclaf) for the gift of RU486, C. Carreaud and S. Sar for technical assistance. MMTV-CAT construction was kindly provided by G. Schütz (Heidelberg) and plasmid PHCwt by A.C.B. Cato (Karlsruhe). The manuscript was typed by V. Francois and N. Malpoint.

Received 12 October; accepted 4 November 1988

1. Groyer, A., Schweizer-Groyer, G., Cadepond, F., Mariller, M. & Baulieu, E. E. *Nature* **328**, 624-626 (1987).
2. Sutherland, R. L. & Murphy, L. C. *Molec. cell. Endocr.* **25**, 5-23 (1982).
3. Richard-Foy, H., Sistare, F. D., Tate Riegel, A., Simons, S. S., Jr. & Hager, G. L. *Molec. Endocr.* **1**, 659-665 (1987).
4. Rajpert, E. J. et al. *J. Steroid Biochem.* **26**, 513-520 (1987).
5. Gluzman, Y. *Cell* **23**, 175-182 (1981).
6. Cato, A. C. B., Miksick, R., Schütz, G., Arnemann, J. & Beato, M. *EMBO J.* **5**, 2237-2240 (1986).
7. Brailly, S. et al. *J. Endocr.* **116**, 427-434 (1988).
8. Moguilewsky, M. & Philibert, D. in *The Antiprogesterin Steroid RU486 and Human Fertility Control* (eds Baulieu E. E. & Segal S. J.) 87-98 (Plenum New York, 1985).
9. Beato, M. *Biochim. Biophys. Acta* **910**, 95-102 (1987).
10. Cato, A. C. B., Skroch, P., Weinmann, J., Butkeraitis, P. & Ponta, H. *EMBO J.* **7**, 1403-1410 (1988).
11. Mullick, A. & Katzenellenbogen, B. S. *Biochem. biophys. Res. Commun.* **135**, 90-97 (1986).
12. Rauch, M., Loosfelt, H., Philibert, D. & Milgrom, E. *Eur. J. Biochem.* **148**, 213-218 (1985).
13. Bailly, A., Le Page, C., Rauch, M. & Milgrom, E. *EMBO J.* **5**, 3235-3241 (1986).
14. Green, S. & Chambon, P. *Nature* **324**, 615-617 (1986).
15. Moguilewsky, M. & Philibert, D. *J. Steroid Biochem.* **20**, 271-276 (1984).
16. Bourgeois, S., Pfahl, M. & Baulieu, E. E. *EMBO J.* **3**, 751-755 (1984).
17. Becker, P. B., Gloss, B., Schmid, W., Strähle, U. & Schütz, G. *Nature* **324**, 686-688 (1986).
18. Rochefort, H. & Borgna, J.-L. *Nature* **292**, 257-259 (1981).
19. Loosfelt, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9045-9049 (1986).
20. Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
21. Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).
22. Logeat, F., Le Cunff, M., Pamphile, R. & Milgrom, E. *Biochem. biophys. Res. Commun.* **131**, 421-427 (1985).
23. Loosfelt, H., Logeat, F., Vu Hai, M. T. & Milgrom, E. *J. biol. Chem.* **259**, 14196-14202 (1984).
24. Logeat, F. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6456-6459 (1983).
25. Herbolme, P., Bourachot, B. & Yaniv, M. *Cell* **39**, 653-662 (1984).
26. Maxam, A. M. & Gilbert, W. C. *Meth. Enzym.* **65**, 499-560 (1980).
27. Logeat, F. et al. *Biochemistry* **24**, 1029-1035 (1985).

Erratum

Timing measurements of the binary millisecond pulsar in the globular cluster M4

J. McKenna & A. G. Lyne
Nature **336**, 226-227 (1988).

THE seventh line of the bold paragraph should read " $\dot{P} = 8.2 \pm 0.4 \times 10^{-19} \text{ s s}^{-1}$ " and not " $\dot{P} = 3.2 \pm 0.4 \times 10^{-19} \text{ s s}^{-1}$ ".

Corrigendum

Temporal integration by a slowly inactivating K^+ current in hippocampal neurons

Johan F. Storm

Nature **336**, 379-381 (1988).

IN this letter, the last two sentences on page 379 as printed are misleading, and should read:

Both the outward current and the underlying conductance were blocked by very low concentrations (30 μM or so) of 4-aminopyridine (4-AP). In this property I_D resembles some other fast-activating, slowly inactivating K^+ currents, for example in axons⁸ and sensory ganglion cells⁹. In contrast, I_D was not blocked by 25 mM tetraethylammonium (TEA), unlike many "delayed rectifier" K^+ currents^{7,8}.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnType ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106